

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup> :</b> C12P 19/18, C07H 15/12 A01H 5/00	<b>A1</b>	<b>(11) International Publication Number:</b> WO 91/19808 <b>(43) International Publication Date:</b> 26 December 1991 (26.12.91)
<b>(21) International Application Number:</b> PCT/US91/04116 <b>(22) International Filing Date:</b> 11 June 1991 (11.06.91)  <b>(30) Priority data:</b> 536,392 11 June 1990 (11.06.90) US  <b>(60) Parent Application or Grant</b> <b>(63) Related by Continuation</b> US 536,392 (CIP) Filed on 11 June 1990 (11.06.90)  <b>(71) Applicant (for all designated States except US):</b> CALGENE, INC. [US/US]; 1920 Fifth Street, Davis, CA 95616 (US).		<b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> STALKER, David, M. [US/US]; 2736 Cumberland Place, Davis, CA 95616 (US). SHEWMAKER, Christine, K. [US/US]; 1409 Spring Creek, Woodland, CA 95616 (US). OAKES, Jannette, V. [US/US]; 2408 Amapola Drive, Davis, CA 95616 (US).  <b>(74) Agents:</b> LASSEN, Elizabeth et al.; Calgene, Inc., 1920 Fifth Street, Davis, CA 95616 (US).  <b>(81) Designated States:</b> AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.  <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> PRODUCTION OF CYCLODEXTRINS IN TRANSGENIC PLANTS  <b>(57) Abstract</b>  The present invention is directed to the production of cyclodextrins in plants. Specifically, it has been found that host plants can be successfully transformed with a DNA sequence capable of expressing cyclodextrin glycosyltransferase which will convert the plants' endogenous starch reserves to cyclodextrins.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CI	Côte d'Ivoire	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark				

### Description

#### Production of cyclodextrins in Transgenic Plants

5

This application is a continuation-in-part of USSN 07/536,392 filed on June 11, 1990.

#### Technical Field

10

This invention relates to the production of cyclodextrins, and, more particularly, to the use of a cyclodextrin glycosyltransferase structural gene to direct the production of cyclodextrins in plants.

#### 15 Background of the Invention

Cyclodextrins are the products of enzymatic starch degradation by a class of amylases termed cyclodextrin glycosyltransferase (CGT) enzymes. The family of cyclodextrins contains three major and several minor cyclic oligosaccharides which are composed of a number of homogenous cyclic  $\alpha$ -1,4-linked glucopyranose units. The cyclodextrin having six glucopyranose units is termed  $\alpha$ -cyclodextrin (also known as Schardinger's  $\alpha$ -dextrin, cyclomaltohexaose, cyclohexaglukan, cyclohexaamylose,  $\alpha$ -CD, ACD and C6A). The seven unit cyclodextrin is termed  $\beta$ -cyclodextrin (also known as Schardinger's  $\beta$ -dextrin, cyclomaltoheptaose, cycloheptaglukan,  $\beta$ -CD, BCD and C7A). The eight unit cyclodextrin is termed  $\gamma$ -cyclodextrin (also known as Schardinger's  $\gamma$ -dextrin, cyclomaltooctaose, cyclooctaglukan, cyclooctaamylose,  $\gamma$ -CD, GCD and C8A).

The cyclic nature of cyclodextrins allows them to function as clathrates (inclusion complexes) in which a guest molecule is enclosed in the hydrophobic cavity of the cyclodextrin host without resort to primary valence forces. Thus, the components are bound as a consequence of geometric factors, and the presence of one component does not significantly affect the structure of the other component. Complexing a hydrophobic compound with

35

cyclodextrin increases the stability and solubility of the hydrophobic compound. Applications of this phenomena have been found in many fields including pharmaceuticals, foods cosmetics and pesticides.

5        In pharmaceutical applications, complexing a drug with cyclodextrins for oral delivery can have many advantages. Among the benefits are the transformation of liquids into solids which can be formed into tablets, stabilization of drugs against volatilization and oxidation, reduction of  
10 bad taste or smell, improvement in the rate of dissolution of poorly soluble drugs and increases in blood levels of poorly water soluble drugs (Pitha, in *Controlled Drug Delivery*, Bruck, ed. Vol. 1, p. 125, (1983) CRC Press). From the limited research done on parenteral administration  
15 of cyclodextrin-complexed drugs, some of the same advantages found for oral delivery can also be observed. The undesirable side effects of drugs can be reduced with complexation with cyclodextrins. Such side affects include gastric irritation from oral delivery, local irritation and  
20 hemorrhagic areas from intramuscular injection, and local irritation from eye-drops (Szejtli, J., *Cyclodextrin Technology*, Kluwer Academic Publications, Boston (1988), pp. 186-306).

      The addition of cyclodextrins to food products or  
25 cosmetics can also have many effects. In spices, food flavoring or perfume fragrances, cyclodextrins protect against oxidation, volatility, and degradation by heat or light (Hashimoto, H., "Application of Cyclodextrins to Food, Toiletries and Other Products in Japan," in  
30 *Proceedings of the Fourth International Symposium of Cyclodextrins*, O. Huber and J. Szejtli, eds. (1988) pp. 533-543). Cyclodextrins can also eliminate or reduce undesirable smells or tastes, and modify food or cosmetic textures.

35        Complexing pesticides with cyclodextrins can increase the bioavailability of poorly wettable or slightly soluble substances, and transform volatile liquids or sublimable solids into stable solid powders (Szejtli, J. (1988) *supra*

at pp. 335-364; U.S. Patent No. 4,923,853). Pesticides which are sensitive to light, heat or oxygen degradation can be stabilized by complexing with cyclodextrins.

Currently, production of cyclodextrins begins with the cultivation of an appropriate microorganism, e.g., *Bacillus macerans*, and separation, purification and concentration of the amylase enzyme. The enzyme is then used to convert a starch substrate to a mixture of cyclic and acyclic dextrins. Subsequent separation and purification of cyclodextrins is then required. The bacterial strain from which the enzyme is isolated and the length of time the starch conversion is allowed to progress determines the predominant form of cyclodextrin produced. Manufacturers of  $\alpha$ -cyclodextrins attempt to manipulate the reaction to preferentially make a specific cyclodextrin, however, the process is not easily controlled and a mixture of cyclodextrins is obtained. At the present time  $\beta$ -cyclodextrin is the most widely commercialized form of cyclodextrins because the  $\beta$ - form is much cheaper to produce than the  $\alpha$ - or  $\gamma$ -cyclodextrins.

In 1987, the U.S. market for cyclodextrins was predicted to reach \$50 million per year within 2 years; that figure would double if the U.S. Food and Drug Administration approved the use of cyclodextrins in food (Seltzer, R., *Chem. Eng. News*, (May 1987) pp. 24-25). The world market is estimated to be twice the U.S. figure (Szejtli, J. (1988) *supra* at p. viii). The potential U.S. market for cyclodextrins has been predicted to reach as high as \$245 million per year (Anon., *Bioproc. Technol.*, Nov. 1987). There is potentially a large market waiting to be tapped if the cost of cyclodextrins could be lowered through alternative production methods.

#### Disclosure of the Invention

Recognizing the disadvantages of bacterial-derived CGT-mediated cyclodextrin production, it is considered desirable to produce cyclodextrins where the CGT is the expression product of a recombinant gene transferred into a

plant host. In this method, generically known as molecular farming, plants are transformed with a structural gene of interest and the product extracted and purified from a harvested field of the transgenic plants. For example,  
5 human serum albumin has been produced in transgenic tobacco and potato (Sijmons, P.C. et al., *Bio/Technology* (1990) 8:217-221).

Extending the idea of molecular farming to cyclodextrins provides a means to lower production costs.  
10 One particularly desirable host plant for such transformation is potato because of the large amount of starch production in potato tubers. A typical tuber contains approximately 16% of its fresh weight as starch (Burton, W.G., *The Potato* (1966) 3rd Edition, Longman  
15 Scientific and Technical Publications, England, p. 361). Transformation of potato plants with the bacterial CGT structural gene linked to a tuber-specific promoter and a leader directing the enzyme, for example, to the amyloplast, provides a means to produce cyclodextrins in  
20 tubers.

It was therefore considered desirable to apply recombinant deoxyribonucleic acid (rDNA) and related technologies to provide for the production of cyclodextrin in modified plants.

25 Proceedings from the seminal work of Cohen & Boyer, U.S. Patent No. 4,237,224, rDNA technology has become available to provide novel DNA sequences and to produce heterologous proteins in transformed cell cultures. In general, the joining of DNA from different organisms relies  
30 on the excision of DNA sequences using restriction endonucleases. These enzymes are used to cut donor DNA at very specific locations, resulting in gene fragments which contain the DNA sequences of interest. Alternatively, structural genes coding for desired peptides and regulatory  
35 control sequences of interest can now be produced synthetically to form such DNA fragments.

These DNA fragments usually contain short single-stranded tails at each end, termed "sticky-ends". These

sticky-ended fragments can then be ligated to complementary fragments in expression vehicles which have been prepared, e.g., by digestion with the same restriction endonucleases. Having created an expression vector which contains the structural gene of interest in proper orientation with the control elements, one can use this vector to transform host cells and express the desired gene product with the cellular machinery available. Recombinant DNA technology provides the opportunity for modifying plants to allow the expression of cyclodextrin glycosyltransferase and the production of cyclodextrins *in vivo*.

However, while the general methods are easy to summarize, the construction of an expression vector containing a desired structural gene is a difficult process and the successful expression of the desired gene product in significant amounts while retaining its biological activity is not readily predictable. Frequently, bacterial-derived gene products are not biologically active when expressed in plant systems.

To successfully modify plants using rDNA, one must usually modify the naturally occurring plant cell in a manner in which the cell can be used to generate a plant which retains the modification. Even in successful cases, it is often essential that the modification be subject to regulation. That is, it is desirable that the particular gene be regulated as to the differentiation of the cells and maturation of the plant tissue. In the case of cyclodextrin glycosyltransferase it is also important that the modification be performed at a site where the product will be directed to contact the starch storage regions of the modified plant. Thus, genetic engineering of plants with rDNA presents substantially increased degrees of difficulty.

In addition, the need to regenerate plants from the modified cells greatly extends the period of time before one can establish the utility of the genetic construct. It is also important to establish that the particular constructs will be useful in a variety of different plant

species. Furthermore, one may wish to localize the expression of the particular construct in specific cell types and it is desirable that the genetically modified plant retain the modification through a number of  
5 generations.

The present invention relates to the production of cyclodextrin in genetically modified plants. In one aspect, the invention comprises a DNA sequence comprising an uninterrupted DNA sequence having a 5'-end and a 3'-end  
10 which codes for the expression of a cyclodextrin glycosyltransferase enzyme together with at least one heterologous DNA sequence bound to either the 5'-end or the 3'-end of said cyclodextrin glycosyltransferase encoding sequence.

15 In accordance with another aspect of the subject invention, a DNA construct is provided which comprises DNA sequences, in the 5' -> 3' direction of transcription, which code for: A transcriptional and translational initiation region functional in a plant cell, and a  
20 structural gene coding for the expression of a cyclodextrin glycosyltransferase enzyme. Optionally, the DNA construct will also contain DNA sequences which code for a transit peptide in reading frame at the 5'-terminus of said cyclodextrin glycosyltransferase encoding sequence, where  
25 the transit peptide is capable of directing transport of the expression product of said cyclodextrin glycosyltransferase encoding sequence to at least one discrete location in a host organism, and/or a transcriptional and translational termination regulatory  
30 region located in the 3' direction from said structural gene.

Additional aspects of this invention provide plant cells containing the present DNA constructs, and methods for using the constructs to produce cyclodextrins in host  
35 plants.



### Brief Description of the Drawings

Figure 1 depicts the DNA sequence (SEQ ID NO: 1) which encodes a SSU transit peptide from soybean plus 48bp of DNA which encodes a mature SSU protein from pea, together with the amino acid sequence (SEQ ID NO: 2) encoded by the reading frame (upper sequence);

Figure 2 depicts a comparison of DNA sequences from patatin 5' untranslated regions from *Solanum tuberosum* varieties Kennebec (top sequence, SEQ ID NO: 3) (generated by PCR) and Maris Piper (bottom sequence, SEQ ID NO: 4);

Figure 3 depicts a comparison of DNA sequences from patatin 5' untranslated regions from *Solanum tuberosum* varieties Russet Burbank (top sequence, SEQ ID NO: 5) (generated by PCR) and Maris Piper (bottom sequence, SEQ ID NO: 4);

Figure 4A depicts a comparison of DNA sequences for native *Klebsiella pneumoniae* cyclodextrin glycosyltransferase (bottom sequence, SEQ ID NO: 6) and PCR-generated pCGT2 cyclodextrin glycosyltransferase (top sequence, SEQ ID NO: 7) (absence of bar between bases indicates difference in the two sequences); and

Figure 4B depicts a comparison of amino acid sequences for native *Klebsiella pneumoniae* cyclodextrin glycosyltransferase (bottom sequence, SEQ ID NO: 8) and pCGT2 cyclodextrin glycosyltransferase (top sequence, SEQ ID NO: 9) (absence of bar between residues indicates difference in the two sequences).

### Detailed Description of the Invention

The present invention is directed to the production of cyclodextrins in plants.

In accordance with one aspect of the subject invention, DNA constructs and methods are provided which permit modification of the composition of host plants to increase synthesis of starch degradation products. It has been found that host plants can be successfully transformed with such DNA constructs which include an amylase structural gene such as the sequence for expression of a

cyclodextrin glycosyltransferase enzyme, to provide for the production of cyclodextrins from endogenous starch reserves in a variety of host plants.

- As used herein, cyclodextrin glycosyltransferase (CGT) is intended to include any equivalent amylase enzyme capable of degrading starch to one or more forms of cyclodextrin. Considerations for use of a specific CGT in plants for the conversion of starch to cyclodextrin include pH optimum of the enzyme and the availability of substrate and cofactors required by the enzyme. The CGT of interest should have kinetic parameters compatible with the biochemical systems found in the host plant cell. For example, the selected CGT may compete for starch substrate with other enzymes.
- The most preferred cyclodextrin forms are the  $\alpha$ -,  $\beta$ - or  $\gamma$  forms, although other higher forms of cyclodextrins, e.g.  $\delta$ -,  $\epsilon$ -,  $\zeta$ - and  $\eta$ - forms, are also possible. Different CGT enzymes produce  $\alpha$ ,  $\beta$ , and  $\gamma$  CDs in different ratios. See, Szejtli, J., *Cyclodextrin Technology* (Kluwer Academic Publications, Boston) (1988), pp. 26-33 and Schmid, G., *TIBTECH* (1989) 7:244-248. In addition, various CGT enzymes can preferentially degrade the starch substrate to favor production of a particular cyclodextrin form. Some CGTs produce primarily  $\beta$ -CDs (Bender, H (1990) *Carb. Res.* 206:257-267; Kimura et al. (1987) *Appl. Microbiol. Biotechnol.* 26:149-153), whereas the *Klebsiella* CGT described in the following examples, produces  $\alpha$ - and  $\beta$ -CDs *in vitro* at a ratio of 20:1 when potato starch is used as the substrate (Bender, H. (1990) *supra*). The use of these different CGTs in transgenic plants could result in different CD profiles and thus different utilities. For example, cyclodextrins have been reported as effective in inhibiting apple juice browning, with  $\beta$ -cyclodextrins producing better results than either  $\alpha$ - or  $\gamma$ -cyclodextrins (*Chemistry and Industry, London* (1988) 13:410). In addition, the inventors in this application have discovered that *in vitro* application of  $\beta$ -CDs to potato tuber slices inhibits discoloration, and *in vitro* application to whole

potato tubers prevents a typical blackspot reaction caused by bruising.

The structural gene for a selected CGT can be derived from cDNA, from chromosomal DNA or may be synthesized, either completely or in part. For example, the desired gene can be obtained by generating a genomic DNA library from a source for CGT, such as a prokaryotic source, e.g. *Bacillus macerans*, *Bacillus subtilis* or, preferably, from *Klebsiella pneumoniae*.

The CGT structural gene can also be derived from a known CGT amino acid sequence in a variety of ways. The gene may be synthesized, complete or in part, particularly where it is desirable to provide plant-preferred codons. Thus, all or a portion of the CGT gene open reading frame may be synthesized using codons preferred by the selected plant host. Plant-preferred codons may be determined, for example, from the codons of highest frequency in the proteins expressed in the largest amount in the selected plant host species.

In general, some or all of the CGT structural gene will be derived from a native gene sequence or from genes substantially homologous to such sequences. However, even in such embodiments it may be desirable to modify all or a portion of the native gene codons, for example to enhance expression, by employing host-preferred codons.

Methods for identifying nucleic acid sequences of interest have found extensive exemplification in the literature, although, in individual situations, different degrees of difficulty may be encountered. Various known techniques include the use of probes where genomic DNA or cDNA libraries may be searched for complementary sequences. In addition, genomic DNA or cDNA may also be used as a template in the polymerase chain reaction (PCR), from which fragments carrying the desired CGT structural gene may be obtained.

To provide for an increased expression of a selected CGT in a host plant, a plant cell is desirably transformed with an expression cassette which includes (in the 5' -> 3'

direction of transcription): (1) A transcriptional and translational initiation region functional in a host plant cell; (2) a structural gene encoding at least one CGT enzyme, and preferably including a sequence encoding a transit peptide in reading frame at the 5'-terminus, where the transit peptide directs transfer of the CGT to the starch-storage region; and (3) a transcriptional and translational termination regulatory region functional in a host plant cell.

In general, as the CGT structural gene is not a plant gene, transcriptional and translational initiation and termination regulatory regions functional in a host plant cell must be provided in order to have expression of the gene in the host plant. The regulatory regions, such as the initiation and termination regions, can be homologous (derived from the original host) or heterologous (derived from a foreign source, or a synthetic sequence) to the plant host.

In the initiation regulatory region, promoters and/or translation initiation signals may be employed, including promoters found in the plant host or other plant species that provide for inducible expression or regulated expression in a plant host. For example, promoter regions may be used from the Ti plasmid T-DNA including the opine synthase transcriptional initiation regions, e.g., the octopine synthase promoter, nopaline synthase promoter, agropine synthase promoters, or the like. Other promoters include viral promoters such as the cauliflower mosaic virus (CaMV) region VI or full-length promoter, the 35S transcriptional initiation region, the promoters and transcriptional initiation region associated with the ribulose-1,5-bisphosphate carboxylase (RuBisCo) genes, e.g., the small sub-unit (SSU), protein genes associated with phaseolin, protein storage, cellulose formation, or the like. Timing of expression, and/or tissue specificity, may be provided by the use of transcriptional regulatory regions having the desired expression specificity. Of particular interest in a presently preferred embodiment of

the invention is a transcriptional initiation region from the patatin gene of potato, which demonstrates preferential expression in the potato tuber, or other promoters which similarly are preferentially expressed in the starch-containing tissue as compared to other plant structures.

A desired promoter region may be identified by the region being 5' from the structural gene in a native configuration, for example, the opine gene, and by restriction mapping and sequencing the promoter may be selected and isolated. Similarly, a desired terminator region may be isolated as the region 3' from the structural gene.

Furthermore, it may be desirable to target the activity of the CGT enzyme to a specific tissue, organelle or region in the host. For example, in potato tubers, starch is stored primarily in the amyloplasts, and thus it is considered desirable to provide a DNA construct which will direct the transport of the expressed CGT to the amyloplast. Transport of the expressed CGT into a particular region of the host may be accomplished by the use of a transit peptide to target a region of interest, such as the amyloplast. The DNA encoding the transit peptide is generally inserted 3' to the promoter sequence(s) and 5' to the CGT structural gene. The transit peptide and processing signal may be derived from any plant protein which is expressed in the cytoplasm and translocated to the region of interest.

A desired transit peptide can be identified by comparing the amino acid sequence encoded by the messenger RNA (mRNA) from the particular protein with the sequence of the mature product. The amino acid sequence encoded by the mRNA beginning at the initiation codon (usually a methionine) and absent from the mature protein will normally be the transit sequence. In addition, fragments from the native transit sequence which retain their transport activity can also be used. A transit peptide of use in the present invention is a sequence capable of directing the translocation of a protein joined to the

transit peptide to the host region of interest and includes the whole native transit peptide, or a functional fragment or mutant thereof. Alternatively, the DNA encoding the transit peptide may be used in combination with DNA  
5 encoding a distinct mature protein, in order to provide a useful cleavage site. This combination may incorporate DNA from the same source, or from two or more different sources, such as a transit peptide from soybean and mature protein from pea. In an embodiment of the present  
10 invention, DNA encoding the transit peptide from the ribulosebisphosphate carboxylase (RuBisCo) small subunit (SSU) protein is used in combination with DNA encoding 16 amino acids of mature small subunit (SSU) protein from pea. Alternatively, a soybean transit peptide may be used in  
15 combination with DNA encoding one amino acid of SSU protein from pea.

A transcriptional and/or translational termination regulatory region may be derived from the 3'-region of the structural gene from which the initiation region was  
20 obtained or from a distinct structural gene. The termination region may be derived from a plant gene or a gene associated with the Ti-plasmid such as the nopaline synthase (nos) termination region.

The various DNA sequences including the CGT structural  
25 gene sequence may be joined together in conventional ways. The sequences may be cloned and joined in the proper orientation to provide for constitutive expression of the structural gene in a plant host.

Methods for synthesizing sequences and bringing the  
30 sequences together are well established in the literature. Where a portion of the structural gene open reading frame is synthesized and a portion is derived from natural sources, the synthesized portion may serve as a bridge  
between two naturally occurring portions, or may provide a  
35 3'-terminus and/or a 5'-terminus. Particularly where the signal sequence and the open reading frame encoding a selected CGT are derived from different genes, synthetic adapters commonly will be employed. In other instances,

polylinkers may be employed, where the various fragments may be inserted at different restriction sites or substituted for a sequence in the polylinker.

In order to join the promoter(s) to the structural gene, the non-coding 5'-region upstream from the structural gene may be removed by endonuclease restriction. Where a convenient restriction site is present near the 5'-terminus of the structural gene, the structural gene may be restricted and an adapter employed for linking the structural gene to a promoter region, where the adaptor provides for lost nucleotides of the structural gene. Alternatively, if no convenient restriction sites are present, the PCR may be used to add sites to either or both ends of the sequences of interest for convenient cloning.

In accordance with one aspect of the invention, host plant cells are transformed with an expression cassette comprising a DNA sequence encoding for at least one CGT enzyme capable of converting starch into oligosaccharides under the regulatory control of promoters capable of directing the expression of a heterologous gene in a plant host cell. The DNA sequence may also include a DNA sequence encoding a transit peptide recognized by the plant host to provide for targeting to a specific region within a tissue of interest.

In developing the expression cassette, the various fragments comprising the regulatory regions and open reading frame may be subjected to different processing conditions, such as ligation, restriction enzyme digestion, resection, *in vitro* mutagenesis, primer repair, use of linkers and adapters, and the like. Thus, nucleotide transitions, transversions, insertions, deletions, or the like, may be performed on the DNA which is employed in the regulatory regions and/or open reading frame. The expression cassette thus may be wholly or partially derived from natural sources, and either wholly or partially derived from sources homologous to the host cell, or heterologous to the host cell. Furthermore, the various DNA constructs (DNA sequences, vectors, plasmids,

expression cassettes) of the invention are isolated and/or purified, or synthesized, and thus are not "naturally occurring." During the construction of the expression cassette, the various fragments of the DNA will usually be  
5 cloned in an appropriate cloning vector, which allows for amplification of the DNA, modification of the DNA or manipulation by joining or removing of sequences, linkers, or the like. Normally, the vectors employed will be capable of replication to at least a relatively high copy  
10 number in an expression system, e.g., in *E. coli*.

Depending upon the manner of introduction of the expression construct into the plant, other DNA sequences may be required. Commonly, the expression cassette will be joined to a replication system functional in prokaryotes,  
15 particularly *E. coli*, so as to allow for cloning of the expression cassette for isolation, sequencing, analysis, and the like. Included with the replication system will usually be one or more markers which may allow for selection in the host; such markers usually involving  
20 biocide resistance, for example antibiotic resistance, heavy metal resistance, cytotoxin resistance, complementation, and the like. Where the DNA construct will be microinjected into the host cell, a marker which allows for selection of those cells in which the injected  
25 DNA has become integrated and functional will usually be desirable. Thus, markers will be selected which can be detected in a plant host.

A number of vectors are readily available for cloning, including such vectors as pBR322, the pUC series, the M13  
30 series, etc. The selected cloning vector(s) will generally have one or more markers which provide for selection of transformants. By appropriate restriction of the vector and cassette, and as appropriate, modification of the ends, by chewing back or filling in overhangs, to provide for  
35 blunt ends, addition of linkers, by tailing, complementary ends can be provided for ligation and joining of the vector to the expression cassette or component thereof.



After each manipulation of the DNA in the development of the cassette, the plasmid will be cloned and isolated and, as required, the particular cassette component analyzed as to its sequence to ensure that the desired  
5 sequence has been obtained, and that the sequences are joined in the proper manner. Depending upon the nature of the manipulation, the desired sequence may be excised from the plasmid and introduced into a different vector or the plasmid may be restricted and the expression cassette  
10 component manipulated, as appropriate. The manner of the transformation of *E. coli* with the various DNA constructs (plasmids and viruses) for cloning is not critical to this invention. Conjugation, transduction, transfection or transformation, for example, calcium chloride or phosphate-mediated transformation, may be variously employed.  
15

The DNA sequence containing the CGT structural gene may then be joined to a wide variety of other DNA sequences for introduction into an appropriate host cell. The companion sequence will depend largely upon the nature of  
20 the host, the manner of introduction of the DNA sequence into the host, and whether episomal maintenance or integration is desired.

Alternatively, temperate viruses may be employed into which the structural gene may be introduced for  
25 introduction into a plant host. Where the structural gene has been obtained from a source having regulatory signals which are not recognized by the plant host, it may be necessary to introduce the appropriate regulatory signals for expression.

Where a virus or plasmid, e.g., tumor inducing plasmid, is employed and has been mapped, a restriction site can be chosen which is downstream from a promoter into which the structural gene may be inserted at the appropriate distance from the promoter. Where the DNA  
30 sequences do not provide an appropriate restriction site, one can digest back portions of the DNA sequence for various times with an exonuclease, such as *Bal31* and insert a synthetic restriction endonuclease site. Methods for  
35

introducing viruses and plasmids into plants are described in the literature (e.g., Matzke and Schulten, J., *Mol. App. Genetics* (1981) 1:39-49).

Of particular interest is the use of a tumor-inducing  
5 plasmid, e.g., Ti or Ri, where the CGT structural gene may be integrated into plant cell chromosomes. By employing the Ti-DNA right and left borders, where the borders flank an insert comprising the CGT structural gene under transcriptional and translational regulatory signals  
10 recognized by the plant host, the construct may be integrated into the plant genome and provided for expression of the CGT in the plant cell at various stages of differentiation.

The constructs of the present invention can be  
15 introduced into a variety of plant hosts in a variety of ways and, for example, may be present as an episomal element or integrated into the host chromosome. For example, the structural gene as part of a construct may be introduced into a plant cell nucleus by micropipet  
20 injection for integration by recombination into the host genome. Transformed plants of this invention include cells which have experienced in vitro addition of DNA as well as progeny carrying the added DNA. By plant cell is meant discrete cells, plant organized or unorganized tissue,  
25 plant parts and whole plants. Plant cells may be transformed in vitro by co-cultivation with *Agrobacterium*, electroporation, protoplast fusion, microinjection, bombardment with microprojectiles and the like.

Plasmids used in plant transformation which may be  
30 transformed into *Agrobacterium tumefaciens* are often called binary vectors. In addition to the transcription regulatory regions, a binary vector may contain the left and more preferably at least a right border of the Ti plasmid from *A. tumefaciens*. The vector may contain  
35 origins of replication active in *E. coli* and *Agrobacterium* so that the plasmid may be replicated in either host. To allow for selection of host cells carrying the binary vector, a selectable marker may be joined to the other

components of the vector, i.e., the DNA construct. This marker is preferably an antibiotic resistance marker such as a gene coding for resistance to gentamicin, chloramphenicol, kanamycin, ampicillin, and the like.

5       The genus *Agrobacterium* includes the species *A. tumefaciens*, which causes crown gall disease in plants, and the species *A. rhizogenes*, which causes hairy root disease in plants. The virulence of *A. tumefaciens* may be attributed to the Ti (tumor-inducing) plasmid, and the  
10       virulence of *A. rhizogenes* attributed to the Ri (root-inducing) plasmid. The Ti and Ri plasmids carry regions called T-DNA (transferred DNA) which becomes integrated into the host plant genome, and from there induce tumor or hairy root formation. Conveniently, these plasmids may be  
15       "disarmed" such that the region between the T-DNA regions, which causes tumor induction or hairy root formation, is removed. Subsequently, DNA sequences of interest may be inserted between the T-DNA regions, such constructs commonly being called "expression constructs". This new  
20       DNA sequence is then integrated into the plant genome, along with the T-DNA, resulting in a plant containing in its genome this DNA sequence of interest.

      The DNA construct including the CGT structural gene may be introduced into a wide variety of plants, both  
25       monocotyledon and dicotyledon, which produces starch. Of special interest is the introduction of such a DNA construct into plants which desirably include substantial amounts of endogenous starch in at least one portion of the plant. Representative examples of such host plants include  
30       plants which have an abundance of starch in the seed, such as corn (e.g. *Zea mays*), cereal grains (e.g. wheat (*Triticum* spp.), rye (*Secale cereale*), triticale (*Triticum aestivum* x *Secale cereale* hybrid), etc.), waxy maize, sorghum (e.g. *Sorghum bicolor*) and rice (e.g. *Oryza*  
35       *sativa*), in the root structures, such as potato (e.g., Irish (*Solanum tuberosum*), Sweet (*Ipomoea batatas*), and yam (*Discorea* spp.)), tapioca (e.g. cassava (*Manihot esculenta*)) and arrowroot (e.g., *Marantaceae* spp.,

*Cycadaceae* spp., *Cannaceae* spp., *Zingiberaceae* spp., etc.), or in the stem, such as sago (e.g. *Palmae* spp., *Cycadales* spp.). Starch is also found in some botanical fruits, including for example tomato, apple, pear, etc. The CGT  
5 gene may be present in cells or plant parts including callus, roots, tubers, propagules, plantlets, seed, seedlings, pollen, or the like.

Once the cells are transformed, transgenic cells may be selected by means of a marker associated with the  
10 expression construct. The expression construct will usually be joined with such a marker to allow for selection of transformed plant cells, as against those cells which are not transformed. As before, the marker will usually provide resistance to an antibiotic, e.g., kanamycin,  
15 gentamicin, hygromycin, and the like, or an herbicide, e.g. glyphosate, which is toxic to plant cells at a moderate concentration.

After transformation, the plant cells may be grown in an appropriate medium. In the case of protoplast  
20 transformations, the cell wall will be allowed to reform under appropriate osmotic conditions. In the case of seeds or embryos, an appropriate germination or callus initiation medium would be employed. For transformation in explants, an appropriate regeneration medium would be used.

25 The callus which results from transformed cells may be introduced into a nutrient medium which provides for the formation of shoots and roots, and the resulting plantlets planted and allowed to grow to seed. During the growth, tissue may be harvested and screened for the presence of  
30 expression products of the expression construct. After growth, the transformed hosts may be collected and replanted. One or more generations may then be grown to establish that the CGT structural gene is inherited in Mendelian fashion.

35 The ability to modify the composition of a host plant offers potential means to alter properties of the plant produce. As used herein, "modify the composition of the plant produce" contemplates the replacement of endogenous

starch with oligosaccharides comprising glucopyranose units. These oligosaccharides, cyclodextrins for example, may then be purified away from the other plant components. For example, by modifying crop plant cells by introducing a functional structural gene expressing a selected CGT, one can provide a wide variety of crops which have the ability to produce cyclodextrins, and desirably such production will be effected without damaging the agronomic characteristics of the host plant. In this manner, substantial economies can be achieved in labor and materials for the production of cyclodextrins, while minimizing the detrimental effects of starch degradation on the host plants.

Preferably, the activity of the gene product will be localized in the starch storage organelles, tissues or regions of the host plant, e.g., the amyloplast of a host potato tuber. The CGT structural gene will manifest its activity by mediating the production of cyclodextrins in at least one portion of the genetically modified host plant or cells thereof.

The following examples serve to illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

## 25 Experimental

In the experimental disclosure which follows, all weights are given in grams (g), milligrams (mg), micrograms ( $\mu$ g) or pounds (lb), all concentrations are given as percent by weight (%), molar (M), millimolar (mM) or micromolar ( $\mu$ M), and all volumes are given in cubic feet ( $\text{ft}^3$ ), liters (l) or milliliters (ml), unless otherwise indicated.

In order to demonstrate the use of CGT compounds in accordance with the present invention, the following examples demonstrate the creation of CGT structural gene DNA constructs and the transfer of such constructs into plant expression systems.

Example 1 - Cloning the GCT Coding Region

This example describes the isolation of the coding region for a cyclodextrin glycosyltransferase (CGT) gene from *Klebsiella pneumoniae* and the engineering of the  
5 coding region for subsequent cloning.

Total genomic DNA is prepared from *Klebsiella pneumoniae* M5A1 (Binder et al., *Gene* (1986) 47:269-277) by growing a 5ml culture in ECLB (Maniatis, T. et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, NY (1982)) overnight at 37°C. The bacteria are pelleted by  
10 centrifugation for 10 minutes at 4500xg, the supernatant is discarded, and the pellet is resuspended in 2.5ml of 10mM Tris, 1mM EDTA buffer. To this suspension is added 500µl of a 5mg/ml Pronase® protease (Calbiochem Brand  
15 Biochemicals; La Jolla, CA) solution and 2ml of 2% lauryl sulfate, sodium salt (Sigma; St. Louis, MO), with gentle mixing and the suspension is incubated at 37°C for 50 minutes. A clear solution indicates that the bacteria have lysed. The solution is then extracted with 5ml phenol,  
20 then 5ml phenol:chloroform:isoamyl alcohol (25:24:1), followed by 5ml chloroform. Nucleic acids are precipitated from the aqueous phase with 1/10 volume of 3M sodium acetate and two volumes of 100% ethanol, and the tube is incubated at room temperature for 1 hour. Nucleic acids  
25 are removed from solution and resuspended in 1ml water. A second ethanol precipitation is performed and the nucleic acids are resuspended in 200µl of 10mM Tris, 1mM EDTA buffer.

Oligonucleotide probes flanking the 2kb cyclodextrin glycosyltransferase (EC 2.4.1.19) gene of *K. pneumoniae* (Bender, H., *Arch. Microbiol.* (1977) 111:271-282) and containing restriction sites for *Bam*HI and *Sal*I are synthesized on an Applied Biosystems 380A DNA synthesizer (Foster City, CA) in accordance with the manufacturer's  
35 instructions. Specifically the probes are:

*Bam*HI

str3: 5'ATATAGGATCCATTAGGACTAGATAATGAAAAGAA 3'  
(SEQ ID NO: 10)

Sal I

str4: 5'AATAAGTCGACTTTTAATTAAAACGAGCCATTCGT 3'  
(SEQ ID NO: 11)

5

The nucleic acid preparation of *K. pneumoneae* is treated with RNase and the DNA is used as a template in a polymerase chain reaction (PCR) with str3 and str4 as primers. A Perkin-Elmer/Cetus (Norwalk, CT) thermal cycler is used with the manufacturer's reagents and in accordance with the manufacturer's instructions. The reaction mixture contains 41.5µl H<sub>2</sub>O, 10µl 10X Reaction buffer, 16µl dNTP's [1.25mM dCTP, dATP, dGTP & dTTP], 5µl str3 (20mM), 5µl str4 (20mM), 22µl total *K. pneumoneae* DNA (0.05µg/µl), and 0.5µl Tag polymerase. The reaction is performed for 15 cycles with melting (denaturation) for 1 minute at 94°C, annealing (hybridization) for 2 minutes at 37°C and chain elongation for 3 minutes at 72°C. The reaction is then performed for an additional 10 cycles with melting for 1 minute at 94°C., annealing or 2 minutes at 37°C and chain elongation at 72°C for 3 minutes 15 seconds initially and increasing the time by 15 seconds each cycle so that the last cycle is 5 minutes 45 seconds.

The resulting PCR product fragments (~2kb) are digested with *Sal*I and *Bam*HI and ligated into a *Sal*I and *Bam*HI digest of pCGN65α3X (see below). Transformed *E. coli* DH5α cells (BRL; Gaithersburg, MD) containing pCGN65α3X are screened on 1% starch plates (ECLB + 1% starch) by flooding with I<sub>2</sub>/KI and evaluating for clearing of starch from around the edge of the colony.

Clone 1 exhibited a good zone of clearing and was digested with *Sph*I and *Sal*I, ligated into *Sph*I- and *Sal*I-digested pUC19 (Norrander et al., *Gene* (1983) 26:101-106) and Yanisch-Perron et al., *Gene* (1985) 33:103-119), yielding the plasmid pCGT2 (~4.5kb). Sequence analysis of pCGT2 (Fig. 4A and SEQ ID NOS: 6-7) showed six single base changes randomly distributed throughout the CGT gene (99.7% homology) which resulted in three amino acid changes (Fig. 4B and SEQ ID NOS: 8-9). Plasmid pCGT2 was digested with

SphI, treated with the Klenow fragment of DNA polymerase I (Klenow fragment) to generate blunt ends and to ligate in a BglIII linker. The resulting plasmid, pCGT4, was sequenced using the Sequenase® DNA sequencing kit (U.S. Biochemical; 5 Cleveland, OH) in accordance with the manufacturer's instructions to confirm the correct reading frame:

HindIII      BamHI  
 5'CCA|AGC|TTG|CG|GAT|CCG|CAG|ACG|ATT (SEQ ID NO: 12)  
 10      lac  $\alpha$  ->      -> CGT ->

#### Construction of pCGN65 $\alpha$ 3X

Plasmid pUC18 (Yanisch-Perron et al., (1985) *supra*) is digested with HaeII to release the lacZ' fragment, treated with Klenow fragment to create blunt ends, and the lacZ'-  
 15 containing fragment is ligated into pCGN565RB-H+X (see below), which has been digested with AccI and SphI, and treated with Klenow fragment, resulting in plasmid pCGN565RB $\alpha$ 3X. In pCGN565RB $\alpha$ 3X, the lac promoter is distal to the T-DNA right border. Both clones are positive for  
 20 lacZ' expression when plated on an appropriate host. Each clone contains coordinates 13990-14273 of the T-DNA right border fragment (Barker et al., *Plant mol. Biol.* (1983) 2:335-350), having deleted the AccI-SphI fragment (coordinates 13800-13989). The 728bp BglIII-XhoI fragment  
 25 of pCGN565RB $\alpha$ 3X, containing the T-DNA right border piece and the lacZ' gene, is cloned into BglIII- and XhoI-digested pCGN65 $\Delta$ KX-S+X to replace the BglIII-XhoI right border fragment of pCGN65 $\Delta$ KX-S+X and create pCGN65 $\alpha$ 3X. The construction of pCGN65 $\alpha$ 3X is described in detail in co-  
 30 pending U.S. application Ser. No. 07/382,176, filed July 19, 1989.

#### Construction of pCGN565RB-H+X

Plasmid pCGN451 includes an octopine cassette which contains approximately 1556bp of the 5' non-coding region  
 35 fused, via an EcoRI linker, to the 3' non-coding region of the octopine synthase gene of pTiA6. The pTi coordinates are 11,207 to 12,823 for the 3' region and 13,643 to 15,208 for the 5' region (Barker et al., (1983) *supra*). Plasmid pCGN451 is digested with HpaI and ligated in the presence



of synthetic *Sph*I linker DNA to generate pCGN55. The *Xho*I-*Sph*I fragment of pCGN55 (coordinates 13800-15208, including the right border of *Agrobacterium tumefaciens* T-DNA (Barker et al., *Gene* (1977) 2:95-113) is cloned into *Sal*I- and

5 *Sph*I-digested pUC19 (Yanisch-Perron et al., (1985) *supra*) to create pCGN60. The 1.4kb *Hind*III-*Bam*HI fragment of pCGN60 is cloned into *Hind*III- and *Bam*HI-digested with pSP64 (Promega, Inc.) to generate pCGN1039. Plasmid pCGN1039 is digested with *Sma*I and *Nru*I (deleting

10 coordinates 14273-15208 (Barker et al., (1977) *supra*) and ligated in the presence of synthetic *Bgl*III linker DNA to create pCGN1039 $\Delta$ NS. The 0.47kb *Eco*RI-*Hind*III fragment of pCGN1039 $\Delta$ NS is cloned into *Eco*RI- and *Hind*III-digested pCGN565 to create pCGN565RB. The *Hind*III site of pCGN565RB

15 is replaced with an *Xho*I site by *Hind*III digestion, treatment with Klenow fragment, and ligation in the presence of synthetic *Xho*I linker DNA to create pCGN565RB-H+X.

## 20 Example 2: Plastid Translocating Sequences

This example describes the preparation of DNA sequences encoding transit peptides for use in the delivery of a CGT gene to starch-containing organelles.

### Construction of SSU + *aroA* Transit Peptide

25 Plasmid pCGN1132 contains a 35S promoter-ribulosebiphosphate carboxylase small subunit (5'-35S-SSU) leader plus 48bp of mature small subunit (SSU) protein from pea *aroA* sequence (the gene locus which encodes 5-enolpyruvyl-3-phosphoshikimate synthetase (EC 2.5.1.19)).

30 It is prepared from pCGN1096, a plasmid containing a hybrid SSU protein gene, which carries DNA encoding mature SSU protein from pea, and *Sst*I and *Eco*RI sites 3' of the coding region (used in the preparation of pCGN1115, a plasmid having a 5'-35S-SSU+48-*aroA*-*tml*-3' sequence, and pCGN1129,

35 a plasmid having a 35S promoter in a chloramphenicol resistance gene (*Cam*<sup>r</sup>) backbone).

Construction of pCGN1096

The *aroA* moiety of pCGN1077 is removed by digestion with *SphI* and *SalI*. In its place is cloned the region coding for the mature pea SSU protein, as an *SphI*-*PstI* fragment, which is then excised with *SphI* and *SalI*. The resulting plasmid, pCGN1094, codes for a hybrid SSU protein having the transit peptide of the soybean clone, and the mature portion of the pea clone and carrier *SstI* and *EcoRI* sites 3' of the coding region. The *HindIII* to *BamHI* region of transposon Tn6 (Jorgensen et al., *Mol. Gen. Genet.* (1979) 177:65) encoding the kanamycin resistance gene ( $Kan^r$ ) is cloned into the same sites of pBR322 (Bolivar et al., *Gene* (1977) 2:95-133) generating pDS7. The *BglIII* site 3' of the  $Kan^r$  gene is digested and filled in with the large fragment of *E. coli* DNA polymerase 1 and deoxy-nucleotides triphosphate. An *SstI* linker is ligated into the blunted site, generating plasmid pCGN1093. Plasmid pPMG34.3 is digested with *SalI*, the site filled in as above and *EcoRI* linkers are ligated into the site resulting in plasmid pCGN1092. The latter plasmid is digested with *SstI* and *SmaI* and the  $Kan^r$  gene excised from pCGN1093 with *SstI* and *SmaI* is ligated in, generating pCGN1095. The  $Kan^r$  and *aroA* genes are excised as a piece from pCGN1095 by digestion with *SstI* and *EcoRI* and inserted into the *SstI* and *EcoRI* sites of pCGN1094, producing pCGN1096. Summarizing, pCGN1096 contains (5' -> 3') the following pertinent features: The SSU gene - a polylinker coding for *PstI*, *SalI*, *SstI*, and *KpnI* - the  $Kan^r$  gene - *SmaI* and *BamHI* restriction sites - the *aroA* gene without the original ATG start codon. The construction of pCGN1096 is also described in detail in co-pending U.S. application Ser. No. 06/097,498, filed September 16, 1987.

Plasmid pCGN1096 is digested to completion with *SalI* and then digested with exonuclease *Bal31* (BRL; Gaithersburg, MD) for 10 minutes, thus deleting a portion of the mature SSU gene. The resulting plasmid is then digested with *SmaI* to eliminate the  $Kan^r$  gene and provide blunt ends, recircularized with T4 DNA ligase and

transformed into *E. coli* LC3 (Comai et al., *Science* (1983) 221:370-371), an *aroA* mutant. DNA isolated from *aroA*<sup>+</sup> and Kan<sup>r</sup> colonies is digested with *Bam*HI and *Sph*I and ligated with *Bam*HI- and *Sph*I-digested M13mp18 (Norrande et al.,  
5 *Gene* (1983) 26:101-106 and Yanisch-Perron et al., *Gene* (1985) 33:103-119) DNA for sequencing. Clone 7 has 48bp of the mature SSU gene remaining (Fig. 1), and the 3' end consists of phe-glu-thr-leu-ser. Clone 7 is transformed into *E. coli* strain 71-18 (Yanisch-Perron et al. (1985)  
10 *supra*) and DNA isolated from transformants is digested with *Sph*I and *Cla*I to remove the 0.65kb fragment containing the 48bp of mature protein and the 5' end of the *aroA* gene. Plasmid pCGN1106 (Comai et al., *J. Biol. Chem.* (1988) 263:15104-15109) is also digested with *Sph*I and *Cla*I and  
15 the 6.8kb isolated vector fragment is ligated with the 0.65kb fragment of clone 7 to yield pCGN1115 (5'-35S-SSU+48-*aroA*-*tml*-3').

The 7.2kb plasmid pCGN1180 (35S-SSU+70-*aroA*-*ocs*3') (Comai et al. (1988) *supra*) and the 25.6kb plasmid pCGN594  
20 (Houck, et al., *Frontiers in Applied Microbiology* (1990) 4:1-17) (LB-Gent<sup>r</sup>-*ocs*5'-Kan<sup>r</sup>-*ocs*3'-RB) (construction of pCGN594 is described in co-pending U.S. application Ser. No. 07/382,802, filed July 19, 1989) are digested with *Hind*III and ligated together to yield the 32.8kb plasmid  
25 pCGN1109 (LB-Gent<sup>r</sup>-35S-SSU+70-*aroA*-*ocs*3'-*ocs*5'-Kan<sup>r</sup>-*ocs*3'-RB).

Plasmid pCGN1109 is digested with *Eco*RI to delete an internal 9.1kb fragment containing the SSU leader plus 70bp of the mature SSU gene, the *aroA* gene and its *ocs*3'  
30 terminator, the Amp<sup>r</sup> backbone from pCGN1180 and *ocs*5'-Kan<sup>r</sup>-*ocs*3' from pCGN594. The *Eco*RI digest of pCGN1109 is then treated with Klenow fragment to blunt the ends, and a *Xho*I linker (dCCTCGAGG) (New England Biolabs Inc.; Beverly, MA) is ligated in, yielding pCGN1125 (LB-35S-RB).

35 Plasmid pCGN1125 is digested with *Hind*III and *Bgl*II to delete the 0.72kb fragment of the 35S promoter. This digest is ligated with *Hind*III- and *Bam*HI-digested Cam<sup>r</sup> vector, pCGN786. pCGN786 is a chloramphenicol resistant

pUC based vector formed by insertion of a synthetic linker containing restriction digest sites *EcoRI*, *SalI*, *BglIII*, *PstI*, *XhoI*, *BamHI*, and *HindIII* into pCGN566 (pCGN566 contains the *EcoRI-HindIII* linker of pUC18 inserted into the *EcoRI-HindIII* sites of pUC13-cm (K. Buckley (1985) Ph.D. thesis, University of California at San Diego). The resulting 3.22kb plasmid, pCGN1128, contains the 35S promoter with a 3' multilinker in a *Cam<sup>r</sup>* backbone.

Plasmid pCGN1128 is digested with *HindIII*, treated with Klenow fragment to blunt the ends and ligated with *BglIII* linkers to yield pCGN1129, thus changing the *HindIII* site located 5' to the 35S promoter into a *BglIII* site.

Plasmid pCGN1115 is digested with *SalI* to removed a 1.6kb fragment containing the SSU leader plus 48bp of the mature SSU gene and the *aroA* gene. An *XhoI* digest of pCGN1129 opened the plasmid 3' to the 35S promoter. Ligation of these two digests yielded the 4.8kb plasmid pCGN1132, containing 5'-35S-SSU leader plus 48bp of mature SSU-*aroA*. Plasmid pCGN1132 is digested with *EcoRI*, treated with Klenow fragment to form blunt ends, and ligated with *SacI* linkers (d(CGAGCTCG) New England Biolabs Inc.; Beverly, MA) to yield pCGN1132S, thus changing the *EcoRI* site 3' to the *aroA* gene to a *SacI* site.

Transit Peptide + Cyclodextrin Glycosyltransferase Gene

Plasmid pCGT4 (See Example 1) and pCGN1132S are digested with *BamHI* and *SalI* and ligated together. The resulting plasmid pCGT5 contains 5'-35S-SSU+48-CGT-3'.

Example 3: Cloning of Patatin Regulatory Regions and Preparation of Patatin-5'-nos-3' Expression Cassettes

This example describes the cloning of patatin-5' regulatory regions from two potato varieties and the preparation of patatin-5'-nos-3' expression cassettes pCGN2143 and pCGN2144. Also provided is the cloning of patatin-3' regulatory regions and the preparation of patatin-5'-patatin-3' expression cassettes pCGN2173 and pCGN2174.

Genomic DNA is isolated from leaves of *Solanum tuberosum* var. Russett Burbank and var. Kennebec as

described in Dellaporta et al., *Plant Mol. Biol. Reporter* (1983) 1(4):19-21, with the following modifications:  
approximately 9g fresh weight of leaf tissue is ground, a polytron grinding is not performed and in the final step  
5 the DNA is dissolved in 300µl of 10mM Tris, 1mM EDTA, pH 8.  
A synthetic oligonucleotide, pat1, containing digestion sites for *NheI*, *PstI* and *XhoI* with 24bp of homology of the 5'-region of a 701bp fragment (coordinates 1611 to 2312) 5' to a class I patatin gene, isolated from *Solanum tuberosum*  
10 var. Maris Piper (Bevan et al., *NAR* (1986) 14:4625-4638) is synthesized (Applied BioSystems 380A DNA synthesizer):

pat1:

15           NheI    PstI    XhoI  
5'CAGCAGGCTAGCTCGCTGCAGCATCTCGAGATTTGTCAAATCAGGCTCAAAGATC3'  
(SEQ ID NO: 13)

A second synthetic oligonucleotide, pat2, containing  
20 digestion sites for *BamHI* and *SpeI* with 25bp of homology to the 3' region of the 701bp piece is also synthesized:

pat2:

25           BamHI   SpeI  
5'ACGACGGGATCCCATACTAGTTTTGCAAATGTTCAAATTGTTTTT3'  
(SEQ ID NO: 14)

Using the genomic potato DNA as a template, and pat1  
30 and pat2 as primers, a polymerase chain reaction (PCR) is performed in a Perkin-Elmer/Cetus thermal cycler with the manufacturer's reagents and in accordance with the manufacturer's instructions. The reaction contains 62.5µl H<sub>2</sub>O, 10µl 10X Reaction buffer, 16µl dNTP's [1.25mM dCTP,  
35 dATP, dGTP & dTTP], 5µl pat1 (20mM), 5µl pat2 (20mM), 1µl potato genomic DNA (3µg/µl), 0.5µl Tag polymerase. The PCR is performed for 25 cycles with melting for 1 minute at 94°C, annealing for 2 minutes at 37°C and chain elongation for 3 minutes at 72°C. The resulting PCR product fragments  
40 (approximately 700bp) are digested with *NheI* and *BamHI*.

Plasmid pCGN1586N (5'-D35S-TMV $\Omega$ '-nos'3'; pCGN1586 (described below) having a *NheI* site 5' to the 35S region) is digested with *NheI* and *BamHI* to delete the D35S- $\Omega$ ' fragment. Ligation of *NheI*-*BamHI* digested pCGN1586N, which  
5 contains the nos-3' region, and the PCR fragments yield a patatin-5'-nos-3' cassette with *SpeI*, *BamHI*, *SalI* and *SstI* restriction sites between the 5' and 3' regions for insertion of a DNA sequence of interest.

The 5' regions of two clones, designated pCGN2143 and  
10 pCGN2144, were sequenced. Plasmid pCGN2143 has a Kennebec patatin-5' region that is 702bp in length and 99.7% homologous to the native sequence (as reported by Bevan (1986) *supra*) (Fig. 2). The 5' region of pCGN2144, from  
15 deletion from coordinate 1971 to coordinate 2040. The remainder of the Russet Burbank clone is 97.0% homologous to the native sequence (as reported by Bevan (1986) *supra*) (Fig. 3).

A synthetic oligonucleotide, pat3S, with 24bp of  
20 homology to the 5' region of a 804bp region 3' to a class I patatin gene (Bevan 5000 to 5804):  
pat3S:

SstI  
25 5'CAGCAGGAGCTCGTACAAGTTGGCGAAACATTATTG3' (SEQ ID NO: 15)  
is synthesized. This oligonucleotide contained a restriction enzyme site for *SstI*. A second oligonucleotide, pat4, with 24bp of homology to the 3' region of the 804bp region was also synthesized:  
30 pat4:

NheI    XhoI    PstI  
5'ACGACGGCTAGCTCGCTCGAGCATCTGCAGTGCATATAAGTTCACATTAATATG3'  
(SEQ ID NO: 16)  
35 It contains digestion sites for the enzymes *NheI*, *XhoI* and *PstI*.

Using Russet Burbank genomic potato DNA as a template, a polymerase chain reaction (PCR) as described above is performed for 25 cycles with melting for 1 minute at 94°C,  
40 annealing for 2 minutes at 42°C and chain elongation for 3

minutes at 72°C. A Perkin-Elmer/Cetus thermal cycler is used with the manufacturer's reagents and in accordance with the manufacturer's instructions. Specifically, the reaction contained 53.5µl H<sub>2</sub>O, 10µl 10X reaction buffer, 16µl dNTP's [1.25mM dCTP, dATP, dGTP & dTTP], 5µl pat3S (20mM), 5µl pat4 (20mM), 10µl genomic potato DNA (3µg/µl), 0.5µl *Taq* polymerase. The resulting approximately 800bp PCR product fragments are digested with *Nhe*I and *Sst*I and ligated into pCGN1586N (see below). Sequencing of one clone, designated pCGN2159, showed that the 3' fragment is 823bp in length and 93.6% homologous to Bevan's reported sequence (Bevan (1986) *supra*).

Cloning of the patatin cassettes pCGN2173 and pCGN2174

A patatin cassette consisting of the 5' patatin region from Kennebec and 3' patatin region from Russet Burbank, identified as pCGN2173, was constructed by a three way ligation of the following fragments: the *Nhe*I to *Sst*I Kennebec 5' patatin fragment of pCGN2143 (see above), the *Sst*I to *Nhe*I Russet Burbank 3' patatin fragment of pCGN2159 and the *Nhe*I to *Nhe*I pUC backbone of pCGN1599.

A second patatin cassette, identified as pCGN2174, was constructed by a three way ligation of the *Nhe*I to *Sst*I Russet Burbank 5' patatin fragment of pCGN2144 (see above), the *Sst*I to *Nhe*I Russet Burbank 3' patatin fragment of pCGN2159 and the *Nhe*I to *Nhe*I pUC backbone of pCGN1599.

Construction of pCGN1586/1586N

Plasmid pCGN2113 (6.1kb) contains a double-35S promoter (D35S) and the *tml*-3' region with multiple cloning sites between them, contained in a pUC-derived plasmid backbone bearing an ampicillin resistance gene (*Amp<sup>r</sup>*). The promoter/*tml* cassette is bordered by multiple restriction sites for easy removal. Plasmid pCGN2113 was digested with *Eco*RI and *Sac*I, deleting the 2.2kb *tml*-3' region. Plasmid pBI221.1 (Jefferson, R.A., *Plant Mol. Biol. Reporter* (1987) 5:387-405) is digested with *Eco*RI and *Sac*I to delete the 0.3kb *nos*-3' region. The digested pCGN2113 and pBI221.1 DNAs are ligated together, and the resultant 4.2kb

recombinant plasmid with the tml-3' of pCGN2113 replaced by nos-3' is designated pCGN1575 (5'-D35S-nos-3').

Plasmid pCGN1575 is digested with *Sph*I and *Xba*I, blunt ends generated by treatment with Klenow fragment, and the  
5 ends are ligated together. In the resulting plasmid, pCGN1577, the *Sph*, *Pst*I, *Sal*I and *Xba*I sites 5' of the D35S promoter are eliminated.

Plasmid pCGN1577 is digested with *Eco*RI, the sticky ends blunted by treatment with Klenow fragment, and  
10 synthetic *Bgl*III linkers (d(pCAGATCTG) New England Biolabs Inc.; Beverly, MA) are ligated in. A total of three *Bgl*III linkers are ligated into the *Eco*RI site creating two *Pst*I sites. The resulting plasmid, termed pCGN1579 (D35S-nos-3'), has a 3' polylinker consisting of 5'-*Eco*RI, *Bgl*III,  
15 *Pst*I, *Bgl*III, *Pst*I, *Bgl*III, *Eco*RI-3'.

A tobacco Mosaic Virus omega' (TMV $\Omega$ ') region (Gallie et al., NAR (1987) 15(21):8693-8711) with *Bgl*III, *Nco*I, *Bam*HI, *Sal*I and *Sac*I restriction sites:

20 *Bgl*III  
5'CAGGAGATCTTATTTTACAACAATTACCAACAACAACAACAACAACAACATTAC  
AATTACTATTTACAATTACACCATGGATCCGTCGACGAGCTC3'  
NcoI BamHI SalI SacI  
(SEQ ID NO: 17)

25 is synthesized on a Applied Biosystems® 380A DNA synthesizer and digested with *Bgl*III and *Sac*I. Plasmid pCGN1577 is digested with *Bam*HI and *Sac*I and the synthetic TMV $\Omega$ ' is ligated in between the 5'-D35S and nos-3' regions. The resulting plasmid is designated pCGN1586 (5'-  
30 D35S-TMV $\Omega$ '-nos-3'). Plasmid pCGN1586N is made by digesting pCGN1586 with *Hind*III and filling in the 5' overhang with Klenow fragment, thus forming a *Nhe*I site 5' to the D35S region.

#### 35 Example 4: Preparation of Patatin-5'-CGT-Nos-3' Binary Vectors

This example describes the construction of binary vectors containing: (1) the patatin-5' region from either *Solanum tuberosum* var. Kennebec or var. Russet Burbank, (2)  
40 DNA encoding a transit peptide from soybean RuBisCo SSU



protein, (3) 48bp of DNA encoding 16 amino acids of mature RuBisCo SSU protein from pea, (4) the CGT coding region from *Klebsiella pneumoniae*, and (5) the nos-3' region.

Plasmid pCGN2143 prepared as described in Example 3 is  
5 digested with *SpeI* and *SstI*, opening the plasmid between the patatin-5' region and nos-3' region. Plasmid pCGT5 (see Example 2) was digested with *XbaI* and *SstI* and ligated with pCGN2143 to yield pCGN2151. Plasmid pCGN2151 consists of 5'-Kennebec patatin-SSU+48-CGT-nos3'. Plasmid pCGN2151  
10 is digested with *PstI* and ligated with *PstI*-digested pCGN1558 (see below). This yields the binary vectors pCGN2160a and pCGN2160b.

In pCGN2160a, the 5'-patatin-SSU+48bp-CGT-nos 3' is inserted into pCGN1558 such that it transcribes in the  
15 opposite direction as the 34S-Kan<sup>r</sup>-tml gene. In pCGN2160b, the 5'-patatin-SSU+48bp-CGT-nos-3' is inserted into pCGN1558 such that it transcribes in the same direction as the 35S-Kan<sup>r</sup>-tml gene.

Plasmid pCGN2144 is digested with *SpeI* and *SstI*,  
20 opening the plasmid between the patatin-5' and nos-3' regions. Plasmid pCGT5 is digested with *XbaI* and *SstI* and ligated with pCGN2144 to yield pCGN2152. Plasmid pCGN2152 consists of 5'-Russet Burbank patatin-SSU+48-CGT-nos3'. Plasmid pCGN2152 is digested with *PstI* and ligated with  
25 pCGN1558 (see below) digested with *PstI*. This yields the binary vectors pCGN2161a and pCGN2161b. In pCGN2161a, the 5'-patatin-SSU+48bp-CGT-nos3' is inserted into pCGN1558 such that it transcribes in the opposite direction as the 35S-Kan<sup>r</sup>-tml gene. In pCGN2161b, the 5'-patatin-SSU+48bp-  
30 CGT-nos-3' is inserted into pCGN1558 such that it transcribes in the same direction as the 35S-Kan<sup>r</sup>-tml gene.

#### Construction of pCGN1558

Plasmid pCGN1558 (McBride and Summerfelt, *Plant Mol. Biol.* (1990) 14(27):269-276) is a binary plant  
35 transformation vector containing the left and right T-DNA borders of *Agrobacterium tumefaciens* octopine Ti-plasmid pTiA6 (Currier and Nester, *J. Bact.* (1976) 126:157-165), the gentamicin resistance gene (Gen<sup>r</sup>) of pPH1JI (Hirsch and

Beringer, *Plasmid* (1984) 12:139-141) an *Agrobacterium rhizogenes* Ri plasmid origin of replication from pLJbB11 (Jouanin et al., *Mol. Gen. Genet.* (1985) 201:370-374), a 35S promoter-Kan<sup>r</sup>-tml-3' region capable of conferring  
5 kanamycin resistance to transformed plants, a ColE1 origin of replication from pBR322 (Bolivar et al. (1977) *supra*) and a lacZ' screenable marker gene from pUC18 (Yanish-Perron et al. (1985) *supra*). The construction of pCGN1558 is described in co-pending U.S. application Ser. No.  
10 07/494,722, filed March 16, 1990.

#### Example 5: Preparation of Transgenic Plants

This example describes the transformation of *Agrobacterium tumefaciens* with a CGT gene DNA construct in  
15 accordance with the present invention and the co-cultivation of such *A. tumefaciens* with plant cells to transform host cells and enable the resultant plants to produce cyclodextrins.

#### Transformation of *Agrobacterium tumefaciens*

20 Cells of *Agrobacterium tumefaciens* strain 2760 (also known as LBA4404, Hoekema et al., *Nature* (1983) 303:179-180) are transformed with binary vectors, such as pCGN2160a, pCGN2160b, pCGN2161a and pCGN2161b (as described in Example 4) using the method of Holsters et al. (*Mol.*  
25 *Gen. Genet.* (1978) 163:181-187). The transformed *A. tumefaciens* are then used in the co-cultivation of plants, in order to transfer the CGT construct into an expression system.

The *Agrobacterium* are grown in AB medium (per liter:  
30 6g K<sub>2</sub>HPO<sub>4</sub>, 2.3g NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O, 2g NH<sub>4</sub>Cl, 3g KCl, 5g glucose, 2.5mg FeSO<sub>4</sub>, 246mg MgSO<sub>4</sub>, 14.7mg CaCl<sub>2</sub>, 15g agar) plus 100µg/l gentamicin sulfate and 100µg/l streptomycin sulfate for 4-5 days. Single colonies are inoculated into 10ml of MG/L broth (per liter: 5g mannitol, 1g L-Glutamic acid or  
35 1.15g sodium glutamate, 0.5g KH<sub>2</sub>PO<sub>4</sub>, 0.10g NaCl, 0.10g MgSO<sub>4</sub>·7H<sub>2</sub>O, 1µg biotin, 5g tryptone, 2.5g yeast extract; adjust pH to 7.0) and are incubated overnight in a shaker at 30°C and 180 rpm. Before co-cultivation, the

*Agrobacterium* culture is centrifuged at 12,000xg for 10 minutes and resuspended in 20ml Murashige and Skoog (MS) medium (#510-1118, Gibco; Grand Island, NY).

Cocultivation with Potato Cells

- 5 Feeder plates are prepared by pipetting 0.5ml of a tobacco suspension culture ( $\sim 10^6$  cells/ml) onto 0.8% agar co-cultivation medium containing MS salts (#510-117, Gibco; Grand Island, NY), 1.0mg/l thiamine-HCl, 0.5mg/l nicotinic acid, 0.5mg/l pyridoxine-HCl, 30g/l sucrose, 5 $\mu$ M zeatin  
10 riboside, 3 $\mu$ M 3-indoleacetyl-DL-aspartic acid, pH 5.9. The feeder plates are prepared one day in advance and incubated at 25°C. A sterile 3mm filter paper disk is placed on top of the tobacco cells after they have grown for one day.

- Tubers of *Solanum tuberosum* var. Russet Burbank and  
15 var. Kennebec between the age of 1 and 6 months post-harvest are peeled and washed in distilled water. All subsequent steps are carried out in a flow hood using sterile techniques. For surface sterilization, tubers are immersed in a solution of 10% commercial bleach (sodium  
20 hypochlorite) with 2 drops of Ivory® liquid soap per 100ml for 10 minutes. Tubers are rinsed six times in sterile distilled water and kept immersed in sterile liquid MS medium (#1118, Gibco; Grand Island, NY) to prevent browning.

- 25 Tuber discs (1-2mm thick) are prepared by cutting columns of potato tuber with a 1cm cork borer and slicing the columns to the desired thickness. Discs are placed into the liquid MS medium culture of the transformed *A. tumefaciens* containing the binary vector of interest  
30 ( $1 \times 10^7$ – $1 \times 10^8$  bacteria/ml) until thoroughly wetted. Excess bacteria are removed by blotting discs on sterile paper towels. The discs are co-cultivated with the bacteria for 48 hours on the feeder plates and then transferred to regeneration medium (co-cultivation medium plus 500mg/l  
35 carbenicillin and 100mg/l kanamycin). In 3 to 4 weeks, shoots develop from the discs.

When shoots are approximately 1cm, they are excised and transferred to a 0.8% agar rooting medium containing MS

salts, 1.0mg/l thiamine-HCl, 0.5mg/l nicotinic acid, 0.5mg/l pyridoxine-HCl, 30g/l sucrose, 200mg/l carbenicillin and 100-200mg/l kanamycin at pH 5.9. Plants are rooted two times with at least one rooting taking place on rooting medium with the higher level of kanamycin (200mg/l). Plants which rooted twice are then confirmed as transformed by performing the NPTII blot activity assays (Radke, S.E. et al., *Theo. Appl. Genet.* (1988) 75:685-694). Plants which are not positive for NPTII activity are discarded.

#### Northern Blot Analysis of Transformed Plants

Total RNA is isolated from 5g of tuber tissue (as described by Logeman et al., *Anal. Biochem.* (1987) 163:16-20). Poly-(A)+RNA is purified over oligo(dT) cellulose (as described by Maniatis et al. (1982) *supra*). RNA denaturing gels are run and blotted (as described by Facciotti et al., *Bio/Technology* (1985) 3:241-246). Equivalent amounts of poly-(A)+RNA are run in each lane. A 1.9kb *Bam*HI fragment of pCGT4 containing the CGT gene is used as a probe in the hybridization. The fragment may be isolated from an agarose gel using the Gene Clean® Kit (Bio 101, Inc.; La Jolla, CA) in accordance with the manufacturer's instructions. Nick-translation and hybridization are performed (as described by Shewmaker et al., *Virology* (1985) 140:281-288 except that washes are at 55°C). The washed blot was autoradiographed on Kodak® X-OMat™ AR X-ray film (Rochester, NY) at -70°C.

An autoradiogram of Russet Burbank potatoes each transformed with one of pCGN2160a, pCGN2161a or pCGN2161b shows bands in each of the transformant sample lanes. The bands are 2.3kb in size, corresponding to the size of CGT message RNA. There was no band present in the lane containing RNA from the untransformed control.

#### Example 6: Recovery of Cyclodextrin From Plants

In this example, the recovery and detection of cyclodextrin in transgenic potato tubers is described.

Rooted plants transformed as described in Example 5 are transplanted from rooting medium to a growth chamber (21°C, 16 hour photoperiod with 250-300 $\mu$ E/m<sup>2</sup>/sec light intensity) in soil prepared as follows: For about 340  
5 gallons, combine 800 lb 20/30 sand (approximately 14 cubic feet), 16 cubic feet Fisons® Canadian Peat Moss, 16 cubic feet #3 vermiculite, and approximately 4.5 lb hydrated lime in a Gleason® mixer. The soil is steamed in the mixer for two hours; the mixer mixes for about 15 seconds at interval  
10 of fifteen minutes over a period of one hour to ensure even heating throughout the soil. During and after the process of steaming, the soil reaches temperatures of at least 180°F for one hour. The soil then sits in the mixer until the next day. At that time, hydrated lime is added, if  
15 necessary, to adjust the pH to range between 6.30 and 6.80.

The relative humidity of the growth chamber is maintained at 70-90% for 2-4 days, after which the humidity is maintained at 40-60%. When plants are well established in the soil, at approximately two weeks, they are  
20 transplanted into the greenhouse. Plants are grown in 6.5 inch pots in a soil mix of peat:perlite:vermiculite (11:1:9) at an average temperature of 24°C day/12°C night. Day length is approximately 12 hours and light intensity levels varied from approximately 600 to 1000 $\mu$ E/m<sup>2</sup>/sec.

25 Tubers are harvested from plants 14 weeks after transplant into the greenhouse. Immediately after harvest, tubers are washed, weighed and their specific gravity determined. Three representative tubers from each transformant are peeled, rinsed in distilled water, chopped  
30 into approximately 0.5 cm cubes, quick frozen in liquid nitrogen, and stored at approximately -70°C until assayed.

#### Extraction of Cyclodextrin

To prepare samples for chromatography, cubes of frozen tuber tissue are ground into a powder in a coffee mill  
35 (Krupps®, Closter, NJ). For each plant assayed, extracts from tubers are prepared as follows: Five grams of frozen potato powder are ground in a prechilled mortar and pestle with 5ml 25% ethanol and then frozen at -70°C for at least

overnight. Samples are then centrifuged at 8500xg for 10 minutes, the supernatant transferred to a clean tube, and the ethanol removed by roto-evaporation for 1 hour.

The cyclodextrin is separated from the tissue samples  
5 in C18 SEP-PAK columns (Waters Chromatography Div.;  
Milford, MA), previously washed with 5ml of 100% methanol,  
followed by 5ml of 50% methanol, followed 5ml of water  
prior to sample application. After the sample is applied,  
the cartridge is washed with 10ml of distilled water to  
10 remove contaminants, and the cyclodextrins are removed with  
0.75ml of 100% methanol, discarding the first two drops.  
The sample is then roto-evaporated to dryness, and  
redissolved in 20 $\mu$ l of 30% methanol.

#### Detection of Cyclodextrin

15 Thin layer chromatography (TLC) is performed as  
described by Szejtli (Szejtli, J., *Cyclodextrin Technology*  
(1988) pp. 20-22, Kluwer Academic Publishers, Boston).  
Samples are spotted on silicagel G plates (#01011,  
Analtech; Newark, DE) and dried. The chromatogram is  
20 developed for approximately 3 hours to a height of 13-15cm,  
with a n-butanol-ethanol-water (4:3:3) mixture. After  
drying, the plate is exposed to iodine vapor for 5-10 min.  
to visualize the chromatogram.

Positive controls of  $\alpha$ -cyclodextrin ( $\alpha$ -CD) and  $\beta$ -  
25 cyclodextrin ( $\beta$ -CD) were run alongside samples from  
transgenic tissue, and average Rf values for four plates  
were 0.39 for  $\alpha$ -CD and 0.36 for  $\beta$ -CD. The  $\alpha$ -CD band  
stained light violet, while the  $\beta$ -CD band stained yellow.  
Tuber tissue from 20 transformed plants was screened for  
30 the presence of  $\alpha$ -CD and  $\beta$ -CD. Tissue of tubers from eight  
Russet Burbank plants (RB2160a-11, RB2160b-7, RB2160b-9,  
RB2161a-2, RB2161b-3, RB2161b-5, RB2161b-11) produced bands  
which stained the same color as the  $\alpha$ -CD control bands and  
had similar Rf values. In addition to the putative  $\alpha$ -CD  
35 bands, the tubers from two plants (RB2160b-7 and 2160b-9)  
produced bands with Rf values and color similar to the  $\beta$ -CD  
control band.

In accordance with one aspect of the subject invention, cyclodextrin can be produced by host plants by incorporation of a cyclodextrin glycosyltransferase structural gene together with the appropriate regulatory sequence. In addition, DNA sequences coding for cyclodextrin glycosyltransferase are provided which can be used for producing cyclodextrin, for example, in methods of the present invention. Thus, plants are grown which can produce cyclodextrin, in order to enhance the utility of the crop plants.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be apparent to those of ordinary skill in the art in light of the teaching of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the claims.

What is claimed is:

1. A DNA sequence comprising an uninterrupted sequence having a 5'-end and a 3'-end which codes for the expression of a cyclodextrin glycosyltransferase enzyme together with at least one heterologous DNA sequence bound to either the 5'-end or the 3'-end of said cyclodextrin glycosyltransferase-encoding sequence.
2. A DNA sequence as recited in Claim 1, wherein said encoding sequence is a cDNA sequence.
3. A DNA sequence as recited in Claim 1, wherein said encoding sequence codes for the expression of a *Klebsiella pneumoniae* cyclodextrin glycosyltransferase enzyme.
4. A DNA sequence as recited in Claim 3, wherein said encoding sequence comprises the *Klebsiella pneumoniae* coding sequence or the synthetic coding sequence equivalent thereof as described in Figure 4A.
5. A DNA sequence as recited in Claim 1, wherein said encoding sequence is bound to a replication system functional in plant host cells.
6. A DNA sequence as recited in Claim 1, wherein the heterologous DNA sequence comprises a sequence coding for a transit peptide capable of directing transport of the expression product of said encoding sequence to at least one discrete location in a host organism.
7. A DNA sequence as recited in Claim 1, further comprising a DNA sequence coding for a marker capable of being identified and selected in a eukaryotic cell containing said sequence.



8. A DNA construct comprising DNA sequences, in the 5' -> 3' direction of transcription, which code for:

a) a transcriptional and translational initiation region functional in a plant cell; and

5 b) a structural gene coding for the expression of a cyclodextrin glycosyltransferase enzyme.

9. A DNA construct as recited in Claim 8, wherein said transcriptional and translational initiation region  
10 comprises at least a portion of a region 5' to a patatin gene from *Solanum tuberosum*.

10. A DNA construct as recited in claim 9, wherein said transcriptional and translational initiation region is  
15 from *Solanum tuberosum* var. Kennebec.

11. A DNA construct as recited in Claim 9, wherein said transcriptional and translational initiation region is from *Solanum tuberosum* var. Russet Burbank.  
20

12. A DNA construct as recited in Claim 8 further comprising:

c) a transcriptional and translational termination regulatory region.  
25

13. A DNA construct as recited in Claim 8 further comprising:

a1) A DNA sequence encoding a transit peptide joined in reading frame at the 5'-terminus of said  
30 cyclodextrin glycosyltransferase-encoding sequence, where the transit peptide is capable of directing transport of the expression product of said cyclodextrin glycosyltransferase-encoding sequence to at least one discrete location in a host organism.

35

14. A DNA construct as recited in Claim 8, further comprising a DNA sequence coding for a marker capable of

being identified and selected in a eukaryotic cell containing said sequence.

15. A plant host cell comprising the DNA construct  
5 of Claim 8 which is capable of expressing at least one functional cyclodextrin glycosyltransferase enzyme.

16. A plant host cell as recited in Claim 15 wherein  
10 the plant is selected from the group consisting of corn, cereal grains, waxy maize, sorghum, rice, potato, tapioca, arrowroot and sago.

17. A plant host cell as recited in Claim 15 wherein  
15 the plant is selected from the group consisting of *Zea mays*, *Triticum* species, *Secale cereale*, *Triticum aestivum* x *Secale cereale* hybrids, *Sorghum bicolor*, *Oryza sativa*, *Solanum tuberosum*, *Ipomoea batatas*, *Discorea* species, *Manihot esculenta*, *Marantaceae* species, *Cycadaceae* species, *Cannaceae* species, *Zingiberaceae* species, *Palmae* species  
20 and *Cycadales* species.

18. A plant which is capable of producing at least one cyclodextrin as a starch degradation product.

25 19. A plant as recited in Claim 18 which comprises the DNA construct of Claim 8, said plant being capable of expressing at least one functional cyclodextrin glycosyltransferase enzyme.

30 20. A plant as recited in Claim 18 wherein the plant is selected from the group consisting of corn, cereal grains, waxy maize, sorghum, rice, potato, tapioca, arrowroot and sago.

35 21. A plant as recited in Claim 18 wherein the plant is selected from the group consisting of *Zea mays*, *Triticum* species, *Secale cereale*, *Triticum aestivum* x *Secale cereale* hybrids, *Sorghum bicolor*, *Oryza sativa*, *Solanum tuberosum*,

*Ipomoea batatas*, *Discorea* species, *Manihot esculenta*,  
*Marantaceae* species, *Cycadaceae* species, *Cannaceae* species,  
*Zingiberaceae* species, *Palmae* species and *Cycadales*  
species.

5

22. A plant as recited in Claim 18 wherein said  
cyclodextrin comprises at least one member selected from  
the group consisting of  $\alpha$ -cyclodextrins,  $\beta$ -cyclodextrins  
and  $\gamma$ -cyclodextrins.

10

23. A method for producing cyclodextrin in plants  
which comprises:

a) modifying at least one plant host cell with  
a DNA construct comprising, in the 5'  $\rightarrow$  3' direction of  
transcription:

i) a transcriptional and translational  
initiation region functional in a plant cell; and  
ii) a structural gene coding for at least  
one cyclodextrin glycosyltransferase enzyme,

20 said DNA sequence being under the transcriptional  
control of said plant host; and

b) maintaining a plant host containing said  
DNA construct under conditions which permit the expression  
of a cyclodextrin-producing amount of cyclodextrin  
glycosyltransferase.

24. The method of Claim 23, wherein said  
transcriptional and translational initiation region  
comprises at least a portion of a region 5' to a patatin  
gene from *Solanum tuberosum*.

25. The method of Claim 24, wherein said  
transcriptional and translational initiation region is from  
*Solanum tuberosum* var. Kennebec.

35

26. The method of Claim 24, wherein said  
transcriptional and translational initiation region is from  
*Solanum tuberosum* var. Russet Burbank.

27. The method of Claim 23 wherein said DNA construct further comprising:

5           ia) a DNA sequence encoding a transit peptide in reading frame at the 5'-terminus of said cyclodextrin glycosyltransferase encoding sequence, where the transit peptide is capable of directing transport of the expression product of said cyclodextrin glycosyltransferase encoding sequence to at least one  
10 discrete location in the plant host organism.

28. The method of Claim 23 where said DNA construct further comprising:

15           iii) a transcriptional and translational termination regulatory region.

29. The method of Claim 23 wherein the plant host is selected from the group consisting of corn, cereal grains, waxy maize, sorghum, rice, potato, tapioca, arrowroot and  
20 sago.

30. The method of Claim 23 wherein the plant host is selected from the group consisting of *Zea mays*, *Triticum* species, *Secale cereale*, *Triticum aestivum* x *Secale cereale*  
25 hybrids, *Sorghum bicolor*, *Oryza sativa*, *Solanum tuberosum*, *Ipomoea batatas*, *Discorea* species, *Manihot esculenta*, *Marantaceae* species, *Cycadaceae* species, *Cannaceae* species, *Zingiberaceae* species, *Palmae* species and *Cycadales* species.

1/21

```

XbaI HindIII EcoRV
| | |
1 TCTAGAAAGCTTGGATATCTGGCAGCAGAAAAACAAGTAGTTGAGAACTAAG 51
SerArgSerLeuAspIleTrpGlnGlnLysAsnLys . LeuArgThrLys
LeuGluAlaTrpIleSerGlySerArgLysThrSerSer . GluLeuAr
. LysLeuGlyTyrLeuAlaAlaGluLysGlnValValGluAsn . G
2 7 16
52 AAGAAGAAATGGCTTCCTCAATGATCTCCTCCCGAGCTGTACCACCGTC 102
LysLysLysMETalaSerSerMETileSerSerProAlaValThrThrVal
gArgArgLysTrpLeuProGln . SerProProGlnLeuLeuProProSe
luGluGluAsnGlyPheLeuAsnAspLeuLeuProSerCysTyrHisArgG
99 99

```

**FIG. 1-1**

2/21 .

	HindII	NaeI	StyI
		HgiCI	NcoI
			DsaI
1103	AACCGTGCCGGTGC	GATGTTGCTC	CATTCAACCGGCTCAAATCCATG
	AsnArgAlaGlyAla	GlyMETValAla	ProPheThrGlyLeuLysSerMET
	rThrValProValPro	AlaTrpLeuLeuHis	SerProAlaSerAsnProTr
	InProCysArgCysArg	HisGlyCysSerIle	HisArgProGlnIleHisG
103		113	150
		118	150
			150

1154 GCTGGCTTCCCCACGAGGAAGACCAACAATGACATTACCTCCATTGCTAGC 204  
AlaGlyPheProThrArgLysThrAsnAsnAspIleThrSerIleAlaSer  
pLeuAlaSerProArgGlyArgProThrMETThrLeuProProLeuLeuAl  
lyTrpLeuProHisGluGluAspGlnGln . HisTyrLeuHisCys . G  
179 200

**FIG. 1-2**

**SUBSTITUTE SHEET**

3/21

SphI  
 BspMI NspI [MfeI]  
 [AvaIII] HaeI  
 Ksp632I  
 | | | | |  
 205 AACGGTGAAGAGTACAATGCATGCAGGTGTGGCCCTCCAATTGGAAGAAG 255  
 AsnGlyGlyArgValGlnCysMETGlnValTrpProProIleGlyLysLys  
 aThrValGluGluTyrAsnAlaCysArgCysGlyLeuGlnLeuGluArgAr  
 lnArgTrpLysSerThrMETHisAlaGlyValAlaSerAsnTrpLysGluG  
 208 222 238  
 220 229 242  
 229

XhoII  
 BamHI  
 |  
 256 AAGTTTGAGACTCTTCTGGGATCC 281  
 LysPheGluThrLeuSerTrpAsp  
 gSerLeuArgLeuPheProGlyIle  
 luVal . AspSerPheLeuGlySer  
 277 277

FIG. 1-3





5/21

[illegible]

**FIG. 2 - 2**

SUBSTITUTE SHEET

	570	580	590	600	610	620
AAACATCAGTAATAGACATCACTAACTTTTATTGGTTATGTCAAACTCAAAATAAAATTCTCAACTTGT						
.....						
AAACATCAGTAATAGACATCACTAACTTTTATTGGTTATGTCAAACTCAAAATAAAATTCTCAACTTGT						

640      650      660      670      680      690  
TTACGTGCCTATATATACCATGCTTGTTATATGCTCAAAGCACCAACAAAATTAAAAACAATTGAACA  
:::~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
TTACGTGCCTATATATACCATGCTTGTTATATGCTCAAAGCACCAACAAAATTAAAAACAATTGAACA

710  
TTTGCAAACTAGTATGGG  
:::~::~:  
TTTGCAAAA  
700 X

6/21

7/21

[illegible]

80	90	100	110	120	130	140
AAAAATAAAGAGG	GTGAGCTAAACA	ATTTCAAATCT	CATCACACAT	TATGGGGT	CAGCCACACAAA	ATAAAA
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
AAAAATAAAGAGG	TTGTGAGCTAAACA	ATTTCAAATCT	CATCACACAT	TATGGGGT	CAGCCACACAAA	ATAAAA

	150	160	170	180	190	200	210
GAACGGTTGGAA	CGGATCTAT	TATATACT	ATAAAGAA	TAGGAAAG	AAAGTGA	GTGAGGT	GCGAG
GGGAGGAGGAG	GGGAGGAG	GGGAGGAG	GGGAGGAG	GGGAGGAG	GGGAGGAG	GGGAGGAG	GGGAGGAG
GAACGGTTGGAA	CGGATCTAT	TATATACT	ATAAAGAA	TAGGAAAG	AAAGTGA	GTGAGGT	GCGAG

[illegible]

**FIG. 3-1**

360

[illegible]

**FIG. 3 - 2**

9/21

	500	510	520	530	540	550
ATAAACATCACGTAATAGACATCACTAACTTTTATTGGTTATGTCAAACCTCAAATAAAAATTTCCTCAACTT						
: :						
ATAAACATCACGTAATAGACATCACTAACTTTTATTGGTTATGTCAAACCTCAAATAAAAATTTCCTCAACTT						

GTTTACGTGCCTATATATACCATGCTTGTTATATGCTCAAAGCACCAACAATAATTAAAAACAATTGAA  
630 640 650 660 670 680 690

GTTTACGTGCCTATATATACCATGCTTGTTATATGCTCAAAGCACCAACAATAATTAAAAACAATTGAA  
630 640 650 660 670 680 690

CATTGCAAACTAGTATGGG  
640 650  
CATTGCAAAA  
700 X

**FIG. 3 - 3**

**SUBSTITUTE CASE**

10/21

X 10 20 30 40 50 60  
GGATCCATTAGGACTAGATAATGAAAAGAACCCGTTTTTTAAATACCTCGGCTGCTATTG  
|||||  
ATTAGGACTAGATAATGAAAAGAACCCGTTTTTTAAATACCTCGGCTGCTATTG  
X 10 20 30 40 50

70 80 90 100 110 120  
CCATTTCGATTGCATTAAATACTTTTTTGTAGCATGCAGACGATTGCTGTGAACCCAG  
|||||  
CCATTTCGATTGCATTAAATACTTTTTTGTAGCATGCAGACGATTGCTGTGAACCCAG  
60 70 80 90 100 110

130 140 150 160 170 180  
AAGAACTTATCTTGATTTTCGTAAGGAGACGATATATTTTCTATTCCCTTGATCGTTTCA  
|||||  
AAGAACTTATCTTGATTTTCGTAAGGAGACGATATATTTTCTATTCCCTTGATCGTTTCA  
120 130 140 150 160 170

190 200 210 220 230 240  
GCGATGGAGATCCAAGTAATAATGCAGGGTTTAATTCGCAACCTACGATCCTAATAATT  
|||||  
GCGATGGAGATCCAAGTAATAATGCAGGGTTTAATTCGCAACCTACGATCCTAATAATT  
180 190 200 210 220 230

FIG. 4A - 1

SUBSTITUTE SHEET

11/21

250 260 270 280 290 300  
TAAAAAATATACTGGAGGAGATCTCCGGGGGTGATTAATAAACTACCCCTATTAAAAAT  
|||||  
TAAAAAATATACTGGAGGAGATCTCCGGGGGTGATTAATAAACTACCCCTATTAAAAAT  
240 250 260 270 280 290

310 320 330 340 350 360  
CACTTGGTGTTACTTCAATCTGGATTAATCTCCCAATCGATAATGTGAATAATACTGATG  
|||||  
CACTTGGTGTTACTTCAATCTGGATTAATCTCCCAATCGATAATGTGAATAATACTGATG  
300 310 320 330 340 350

370 380 390 400 410 420  
CTGCTGGCAATACTGGATATCATGGTTATTGGGGAAGAGATTATTTTCGTATAGATGAAC  
|||||  
CTGCTGGCAATACTGGATATCATGGTTATTGGGGAAGAGATTATTTTCGTATAGATGAAC  
360 370 380 390 400 410

430 440 450 460 470 480  
ATTTTGGCAATCTCGATGATTTCAAAGAAGAACTGACTAGTTTGATGCATAGTCCCTGATTATA  
|||||  
ATTTTGGCAATCTCGATGATTTCAAAGAAGAACTGACTAGTTTGATGCATAGTCCCTGATTATA  
420 430 440 450 460 470

FIG. 4A - 2

SEQUENCE INFORMATION

12/21

```
490      500      510      520      530      540
ATATGAAACTGGTTCTTGATTATGCCCTAATCATTCGAATGCTAATGATGAAAATGAAT
|||||
ATATGAAACTGGTTCTTGATTATGCCCTAATCATTCGAATGCTAATGATGAAAATGAAT
480      490      500      510      520      530

550      560      570      580      590      600
TTGGTGCACTATATCGTGATGGTGTGTTTATTACTGATTATCCTACAGATGTTGCCGCCA
|||||
TTGGTGCACTATATCGTGATGGTGTGTTTATTACTGATTATCCTACGAATGTTGCCGCCA
540      550      560      570      580      590

610      620      630      640      650      660
ATACGGGCTGGTATCATCACAAATGGTGGGTAACGAACTGGAATGATTCTTCCAAGTGA
|||||
ATACGGGCTGGTATCATCACAAATGGTGGGTAACGAACTGGAATGATTCTTCCAAGTGA
600      610      620      630      640      650

670      680      690      700      710      720
AGAAATCATAATCTATTCAATCTATCAGACCTCAATCAATCCAATACTGATGTCTACCAGT
|||||
AGAAATCATAATCTATTCAATCTATCAGACCTCAATCAATCCAATACTGATGTCTACCAGT
660      670      680      690      700      710
```

FIG. 4H - 3



13/21

```
730      740      750      760      770      780
ACTTGTGGATGGCTCTAAATTTGGATCGATGCTGGTGGATGCTATCAGGATTGATG
|||||
ACTTGTGGATGGTCTCTAAATTTGGATCGATGCTGGTGGATGCTATCAGGATTGATG
720      730      740      750      760      770

790      800      810      820      830      840
CCATCAAGCATATGGACAAGTCTTTTATACAGAAATGGACCAGCGATATTTATGATTACA
|||||
CCATCAAGCATATGGACAAGTCTTTTATACAGAAATGGACCAGCGATATTTATGATTACA
780      790      800      810      820      830

850      860      870      880      890      900
GTAAGTCTATCGGCCCGGGAAGGATTTTTTTCTTCGGTGAATGGTTGGTGCCAGTGCGA
|||||
GTAAGTCTATCGGCCCGGGAAGGATTTTTTTCTTCGGTGAATGGTTGGTGCCAGTGCGA
840      850      860      870      880      890

910      920      930      940      950      960
ATACTACAACAGGTGTTGATGGTAATGCTATCGATTACGCCCAACACTTCCGGGTCAGCGT
|||||
ATACTACAACAGGTGTTGATGGTAATGCTATCGATTACGCCCAACACTTCCGGGTCAGCGT
900      910      920      930      940      950
```

FIG. 4A - 4

14/21

970 980 990 1000 1010 1020  
TGCTGGATTTTGGATTCGCGGATACCTTAGAAAGAGTTTGGTAGGACGTAGCGGAAATA  
|||||  
TGCTGGATTTTGGATTCGCGGATACCTTAGAAAGAGTTTGGTAGGACGTAGCGGAAATA  
960 970 980 990 1000 1010

1030 1040 1050 1060 1070 1080  
CAATGAAAACGTTAAATAGTTATCTGATAAAAGACAAACAGTCTTTACCAGTGATGACT  
|||||  
CAATGAAAACGTTAAATAGTTATCTGATAAAAGACAAACAGTCTTTACCAGTGATGACT  
1020 1030 1040 1050 1060 1070

1090 1100 1110 1120 1130 1140  
GGCAGGTTGTTTTATGGATAACCATGATATGGCACGCATTGGTACCGCTCTGCGTTCAA  
|||||  
GGCAGGTTGTTTTATGGATAACCATGATATGGCACGCATTGGTACCGCTCTGCGTTCAA  
1080 1090 1100 1110 1120 1130

1150 1160 1170 1180 1190 1200  
ACGCCACTACTTTTGGTCCTGGAAATAATGAAACCGGTGGAAGTCAGAGTGAAGCTTTTG  
|||||  
ACGCCACTACTTTTGGTCCTGGAAATAATGAAACCGGTGGAAGTCAGAGTGAAGCTTTTG  
1140 1150 1160 1170 1180 1190

FIG. 4A - 5

SUBSTITUTE SHEET

15/21

1210 1220 1230 1240 1250 1260  
CTCAGAAACGTATAGACCTCGGTCGGTTGGGACAATGACTGTACGTGGTATTCCTGCCA  
|||||  
CTCAGAAACGTATAGACCTCGGTCGGTTGGGACAATGACTGTACGTGGTATTCCTGCCA  
1200 1210 1220 1230 1240 1250

1270 1280 1290 1300 1310 1320  
TTTATTATGGTACTGAACATTATGCCGCTAACTTTACCTCTAACAGTTTGGTCAAGTTG  
|||||  
TTTATTATGGTACTGAACATTATGCCGCTAACTTTACCTCTAACAGTTTGGTCAAGTTG  
1260 1270 1280 1290 1300 1310

1330 1340 1350 1360 1370 1380  
GCAGTGATCCTTACAACCGAGAGAAAAATGCCAGGATTTGATACGGAAAGTGAGGCTTCT  
|||||  
GCAGTGATCCTTACAACCGAGAGAAAAATGCCAGGATTTGATACGGAAAGTGAGGCTTCT  
1320 1330 1340 1350 1360 1370

1390 1400 1410 1420 1430 1440  
CCATTATTAAAAACACTGGGTGACCTAAGGAAAAAGTAGCCCGCAATTCAAAATGGAACCT  
|||||  
CCATTATTAAAAACACTGGGTGACCTAAGGAAAAAGTAGCCCGCAATTCAAAATGGAACCT  
1380 1390 1400 1410 1420 1430

FIG. 4H - 6

16/21

1450 1460 1470 1480 1490 1500  
ATACTGAACCTATGGGTTAATGATGATATATAGTATTGAGCGCGCTTCTGGGAACGATA  
| | | | |  
ATACTGAACCTATGGGTTAATGATGATATATAGTATTGAGCGCGCTTCTGGGAACGATA  
1440 1450 1460 1470 1480 1490

1510 1520 1530 1540 1550 1560  
TTGTTATTGTTGCACTTAATCGTGGTGAGGCTAACACAATTAATGTTAAAAATATAGCGG  
| | | | |  
TTGTTATTGTTGCACTTAATCGTGGTGAGGCTAACACAATTAATGTTAAAAATATAGCGG  
1500 1510 1520 1530 1540 1550

1570 1580 1590 1600 1610 1620  
TTCCTAATGGGGTATATCCGAGTTTGATTGGGAATAATAAGTGTTCAGTAGCAATAAAT  
| | | | |  
TTCCTAATGGGGTATATCCGAGTTTGATTGGGAATAATAAGTGTTCAGTAGCAATAAAT  
1560 1570 1580 1590 1600 1610

1630 1640 1650 1660 1670 1680  
AGGCAACACTAACACTTATGCAAAATGAAGCTGTTGTCAATTCGCTCACAAATCAGATGATG  
| | | | |  
GGACAACACTAACACTTATGCAAAATGAAGCTGTTGTCAATTCGCTCACAAATCAGATGATG  
1620 1630 1640 1650 1660 1670

FIG. 4A - 7

17/21

1690 1700 1710 1720 1730 1740  
CGGAGAACCCTACAGTACAAAGCATAAACTTCGCATGTAATAACGGTTATACGATTTCAG  
|||||  
CGGAGAACCCTACAGTACAAAGCATAAACTTCACATGTAATAACGGTTATACGATTTCAG  
1680 1690 1700 1710 1720 1730

1750 1760 1770 1780 1790 1800  
GTCAAAGTGTTTATATTGGTAATATACCTCAGTTAGGTGGTGGACTTAACTAAAG  
|||||  
GTCAAAGTGTTTATATTGGTAATATACCTCAGTTAGGTGGTGGACTTAACTAAAG  
1740 1750 1760 1770 1780 1790

1810 1820 1830 1840 1850 1860  
CGGTAAAAATATCACCACACAAATATCCACAATGGAGTGGAGCTTAGAGCTTCCTTCTG  
|||||  
CGGTAAAAATATCACCACACAAATATCCACAATGGAGTGGAGCTTAGAGCTTCCTTCTG  
1800 1810 1820 1830 1840 1850

1870 1880 1890 1900 1910 1920  
ACTTAAATGTTGAATGGAAGTGTGAAACGTAATGAAACCAATCCGACGGCTAATGTTG  
|||||  
ACTTAAATGTTGAATGGAAGTGTGAAACGTAATGAAACCAATCCGACGGCTAATGTTG  
1860 1870 1880 1890 1900 1910

FIG. 4B - B

18/21

1930	1940	1950	1960	1970	1980
AGTGGCAGTCTGGTGCAAAATAACCAAGTTCAATAGCAATGACACACACAAACGAATGGCT					
AGTGGCAGTCTGGTGCAAAATAACCAAGTTCAATAGCAATGACACACACAAACGAATGGCT					
1920	1930	1940	1950	1960	1970

1990	X
CGTTTAAATTAAAGTCGAC	
CGTTTAAATTAAAA	
1980	X

FIG. 4R - 9

19/21

X 10 20 30 40 50 60  
 MKRNRFFNTSAAIAISIALNTFFCSMQTIAAEPEETYLDFRKETIYFLDFRSDGDP  
 SN  
 |||||  
 MKRNRFFNTSAAIAISIALNTFFCSMQTIAAEPEETYLDFRKETIYFLDFRSDGDP  
 SN  
 X 10 20 30 40 50 60

70 80 90 100 110 120  
 NAGENSATYDPNNLKKYTGGLRGLINKLPYLKSLGVTSIWITPPIDNVNNTDAA  
 GNTGY  
 |||||  
 NAGENSATYDPNNLKKYTGGLRGLINKLPYLKSLGVTSIWITPPIDNVNNTDAA  
 GNTGY  
 70 80 90 100 110 120

130 140 150 160 170 180  
 HGYWGRDYFRIDEHFGNLDDEKELTSLMHSPPDYNMKLVLDYAPNHSNANDENE  
 FGALYRD  
 |||||  
 HGYWGRDYFRIDEHFGNLDDEKELTSLMHSPPDYNMKLVLDYAPNHSNANDENE  
 FGALYRD  
 130 140 150 160 170 180

190 200 210 220 230 240  
 GVFI TDYPTDVAANTGWYHHNGGVTNNWDDFFQVKNNHNLSDLNQSN  
 TDVYQYLLDGSK  
 |||||  
 GVFI TDYPTDVAANTGWYHHNGGVTNNWDDFFQVKNNHNLSDLNQSN  
 TDVYQYLLDGSK  
 190 200 210 220 230 240

FIG. 4B - 1

20/21

250 260 270 280 290 300  
FWIDAGVDAIRDAIKHMDKSFQKWTSDIYDYSKISGREGFFFGFWFGASANTTTGVD  
|||||  
FWIDAGVDAIRDAIKHMDKSFQKWTSDIYDYSKISGREGFFFGFWFGASANTTTGVD  
250 260 270 280 290 300

310 320 330 340 350 360  
GNAIDYANTSGSALLDFGFRDTLERVLVGRSGNTMKTLSYLIKRTVFTSDDWQVVFMD  
|||||  
GNAIDYANTSGSALLDFGFRDTLERVLVGRSGNTMKTLSYLIKRTVFTSDDWQVVFMD  
310 320 330 340 350 360

370 380 390 400 410 420  
NHDMARIGTALRSNATTFGPGNNETGGSQSEAFQAKRIDLGLVATMTVRGIPAIYYGTEH  
|||||  
NHDMARIGTALRSNATTFGPGNNETGGSQSEAFQAKRIDLGLVATMTVRGIPAIYYGTEH  
370 380 390 400 410 420

430 440 450 460 470 480  
YAANFTSNFSGQVGSDPYNREKMPGFDTESEAFSIKTLGDLRKSSPAIQNGTYTELWVN  
|||||  
YAANFTSNFSGQVGSDPYNREKMPGFDTESEAFSIKTLGDLRKSSPAIQNGTYTELWVN  
430 440 450 460 470 480

FIG. 4B - 2

© 1991 by Genentech, Inc.



21/21

490 500 510 520 530 540  
DDILVFERRSGNDIVIVALNRGEANTINVKNIAVPNGVYPSLIGNNSVSVANKQATLTLM  
|||||  
DDILVFERRSGNDIVIVALNRGEANTINVKNIAVPNGVYPSLIGNNSVSVANKRTTTLTLM  
490 500 510 520 530 540

550 560 570 580 590 600  
QNEAVVIRSQSDDAENPTVQSINFACNNGYTTISGQSVYIIGNIPQLGGWDLTKAVKISPT  
|||||  
QNEAVVIRSQSDDAENPTVQSINFACNNGYTTISGQSVYIIGNIPQLGGWDLTKAVKISPT  
550 560 570 580 590 600

610 620 630 640 650 X  
QYPQWSASLELPSDLNVEWKCVKRNETNPTANVEWQSGANNQFNSNDTQTNGSF.  
|||||  
QYPQWSASLELPSDLNVEWKCVKRNETNPTANVEWQSGANNQFNSNDTQTNGSF.  
610 620 630 640 650 X

FIG. 4B - 3

SUBSTITUTE SHEET

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/04116

<b>I. CLASSIFICATION OF SUBJECT MATTER</b> (In several classification symbols apply, indicate all) * According to International Patent Classification (IPC) or to both National Classification and IPC IPC(5): C12P 19/18; C07H 15/12; A01H 5/00 U.S. CL.: 435/97; 536/27; 800/200, 205														
<b>II. FIELDS SEARCHED</b> <div style="text-align: center; border-top: 1px solid black; border-bottom: 1px solid black; margin: 5px 0;">Minimum Documentation Searched †</div> <table style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 20%; border-bottom: 1px solid black;">Classification System</th> <th style="border-bottom: 1px solid black;">Classification Symbols</th> </tr> <tr> <td style="vertical-align: top; padding: 5px;">U.S.</td> <td style="padding: 5px;">435/69.1, 97, 172.3, 240.4 536/27 800/200, 205</td> </tr> </table> <div style="text-align: center; border-top: 1px solid black; border-bottom: 1px solid black; margin: 5px 0;">Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched ‡</div> <p style="margin-top: 10px;">USPTO Automated Patent System, DIALOG files, BIOTECH See attachment for search terms.</p>			Classification System	Classification Symbols	U.S.	435/69.1, 97, 172.3, 240.4 536/27 800/200, 205								
Classification System	Classification Symbols													
U.S.	435/69.1, 97, 172.3, 240.4 536/27 800/200, 205													
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT †</b> <table style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 10%; border-bottom: 1px solid black;">Category *</th> <th style="border-bottom: 1px solid black;">Citation of Document, †† with indication, where appropriate, of the relevant passages ‡‡</th> <th style="width: 10%; border-bottom: 1px solid black;">Relevant to Claim No. ‡‡</th> </tr> <tr> <td style="vertical-align: top; text-align: center; padding: 5px;">Y</td> <td style="padding: 5px;">Plant Molecular Biology, Volume 12, Issued January 1989, WENZLER ET AL., "Analysis of a chemeric class-I patatin-GUS gene in transgenic potato plants: High level expression in tubers and sucrose-inducible expression in cultured leaf and stem explants", pages 41-50. See entire document.</td> <td style="vertical-align: top; text-align: center; padding: 5px;">1-30</td> </tr> <tr> <td style="vertical-align: top; text-align: center; padding: 5px;">Y</td> <td style="padding: 5px;">Science, Volume 229, Issued 20 September 1985, KNORR ET AL., "Biotechnology in Food Production and Processing", pages 1224-1229. See entire document.</td> <td style="vertical-align: top; text-align: center; padding: 5px;">1-30</td> </tr> <tr> <td colspan="3" style="text-align: center; padding: 20px;">(cont. on extra sheet)</td> </tr> </table>			Category *	Citation of Document, †† with indication, where appropriate, of the relevant passages ‡‡	Relevant to Claim No. ‡‡	Y	Plant Molecular Biology, Volume 12, Issued January 1989, WENZLER ET AL., "Analysis of a chemeric class-I patatin-GUS gene in transgenic potato plants: High level expression in tubers and sucrose-inducible expression in cultured leaf and stem explants", pages 41-50. See entire document.	1-30	Y	Science, Volume 229, Issued 20 September 1985, KNORR ET AL., "Biotechnology in Food Production and Processing", pages 1224-1229. See entire document.	1-30	(cont. on extra sheet)		
Category *	Citation of Document, †† with indication, where appropriate, of the relevant passages ‡‡	Relevant to Claim No. ‡‡												
Y	Plant Molecular Biology, Volume 12, Issued January 1989, WENZLER ET AL., "Analysis of a chemeric class-I patatin-GUS gene in transgenic potato plants: High level expression in tubers and sucrose-inducible expression in cultured leaf and stem explants", pages 41-50. See entire document.	1-30												
Y	Science, Volume 229, Issued 20 September 1985, KNORR ET AL., "Biotechnology in Food Production and Processing", pages 1224-1229. See entire document.	1-30												
(cont. on extra sheet)														
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>* Special categories of cited documents: ‡‡</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="width: 45%;"> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"Δ" document member of the same patent family</p> </div> </div>														
<b>IV. CERTIFICATION</b> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; border-bottom: 1px solid black; padding: 5px;">           Date of the Actual Completion of the International Search—  <div style="text-align: center; margin-top: 10px;">19 September 1991</div> </td> <td style="width: 50%; border-bottom: 1px solid black; padding: 5px;">           Date of Mailing of this International Search Report  <div style="text-align: center; margin-top: 10px; font-size: 1.2em;">24 OCT 1991</div> </td> </tr> <tr> <td style="border-bottom: 1px solid black; padding: 5px;">           International Searching Authority  <div style="text-align: center; margin-top: 10px;">ISA/US</div> </td> <td style="border-bottom: 1px solid black; padding: 5px;">           Signature of Authorized Officer <i>P Rhodes</i>  <div style="text-align: center; margin-top: 10px;">P. Rhodes, Examiner (vsh)</div> </td> </tr> </table>			Date of the Actual Completion of the International Search— <div style="text-align: center; margin-top: 10px;">19 September 1991</div>	Date of Mailing of this International Search Report <div style="text-align: center; margin-top: 10px; font-size: 1.2em;">24 OCT 1991</div>	International Searching Authority <div style="text-align: center; margin-top: 10px;">ISA/US</div>	Signature of Authorized Officer <i>P Rhodes</i> <div style="text-align: center; margin-top: 10px;">P. Rhodes, Examiner (vsh)</div>								
Date of the Actual Completion of the International Search— <div style="text-align: center; margin-top: 10px;">19 September 1991</div>	Date of Mailing of this International Search Report <div style="text-align: center; margin-top: 10px; font-size: 1.2em;">24 OCT 1991</div>													
International Searching Authority <div style="text-align: center; margin-top: 10px;">ISA/US</div>	Signature of Authorized Officer <i>P Rhodes</i> <div style="text-align: center; margin-top: 10px;">P. Rhodes, Examiner (vsh)</div>													

III DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category	Citation of Document, with indication where appropriate, of the relevant passages	Relevant to Claim No.
Y	Gene, Volume 47, Issued 1986, Binder et al, "Cyclodextrin-glycosyltransferase from <u>Klebsiella pneumoniae</u> M5a1: cloning, nucleotide sequence and expression", pages 269-277. See entire document.	1-30
Y	Journal of General Microbiology, Volume 134, Issued January 1988, Kaneko et al, "Molecular cloning and nucleotide sequence of the cyclomaltodextrin glucanotransferase gene from the alkalophilic <u>Bacillus</u> sp. Strain No. 38-2", pages 97-105. See entire document.	1-30
Y	Journal of Bacteriology, Volume 169, Number 9, Issued September 1987, Kimura et al, "Nucleotide sequence of the $\beta$ -cyclodextrin glucanotransferase gene of alkalophilic <u>Bacillus</u> sp. Strain 1011 and similarity of its amino acid sequence to those of $\alpha$ -amylases", pages 4399-4402. See entire document.	1-30
Y	Journal of Bacteriology, Volume 166, Number 3, Issued June 1986, Takano et al, "Molecular cloning, DNA nucleotide sequencing, and expression in <u>Bacillus subtilis</u> cells of the <u>Bacillus macerans</u> cyclodextrin glucanotransferase gene", pages 1118-1122. See entire document.	1-30
Y	Agricultural Biological Chemistry, Volume 50, Number 8, Issued August 1986, Kato et al, "Cloning and expression of the <u>Bacillus subtilis</u> No. 313 $\gamma$ -cyclodextrin forming CGTase gene in <u>Escherichia coli</u> ", pages 2161-2162. See entire document.	1-30
Y	Agricultural Biological Chemistry, Volume 51, Number 7, Issued July 1987, Hamamoto et al, "Nucleotide sequence of the cyclomaltodextrin glucanotransferase (CGTase) gene from alkalophilic <u>Bacillus</u> sp. Strain No. 38-2", pages 2019-2022. See entire document.	1-30
Y	Food Technology, Volume 42, Issued January 1988, Pszczola, "Production and potential food applications of cyclodextrins", pages 96-100. See entire document.	1-30
(cont.)		

## FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

X	EP, A, 0, 220, 714 (BOCK et al) 06 May 1987. See entire document.	1-4,6
Y	US, A, 4, 801, 540 (HIATT et al) 31 January 1989. See entire document.	1-30
Y	US, A, 4, 771, 002 (GELVIN) 13 September 1988. See entire document.	1-30

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE<sup>1</sup>

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers \_\_\_\_\_, because they relate to subject matter <sup>12</sup> not required to be searched by this Authority, namely:
  
2. ☐ Claim numbers \_\_\_\_\_, because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out <sup>13</sup>, specifically:
  
3. ☐ Claim numbers \_\_\_\_\_, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING<sup>2</sup>

This International Searching Authority found multiple inventions in this international application as follows:

See attachment.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.
2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:
3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:
4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

## Remark on Protest

- ☐ The additional search fees were accompanied by applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

Attachment to Form PCT/ISA/210:

Part VI. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

Group I. Claims 1-17 and 19-30, drawn to a DNA construct and the cells and plants containing same and to a process of use to make cyclodextrin, classified in Classes 435, 536, and 800, subclasses 97, 27, and 205, respectively, for example.

Group II. Claim 18, drawn to a plant capable of making cyclodextrin, classified in Class 800, subclass 200, for example.

The plant product of Group II does not require the particulars of the DNA construct of Group I as claimed. The plant of Group II may occur naturally or by mutagenesis, for example. Since these two products can be separately made and can have different properties as claimed, these are separate and distinct inventions that are not linked into one inventive concept.